

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd

Gen nucleic nucleic search, using sw model

Run on: October 30, 2003, 21:37:02 : Search time 1194 Seconds
without alignments
11157.539 Million cell updates/sec

File: US 10-028-374-1

Perfect score: 4931

Sequence: 1 cccgcgtccgacatggggc.....cccccaaaaaaaaaaaaaaaaaa 4931

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Database: 2552756 seqs, 134971917 residues

Total number of hits satisfying chosen parameters: 516512

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: 1

N Geneseq 1900003:
1: /SIDSI/gcgdata/geneseq/geneseq-emb/NA1990.DAT:
2: /SIDSI/gcgdata/geneseq/geneseq-emb/NA1981.DAT:
3: /SIDSI/gcgdata/geneseq/geneseq-emb/NA1982.DAT:
4: /SIDSI/gcgdata/geneseq/geneseq-emb/NA1983.DAT:
5: /SIDSI/gcgdata/geneseq/geneseq-emb/NA1984.DAT:
6: /SIDSI/gcgdata/geneseq/geneseq-emb/NA1985.DAT:
7: /SIDSI/gcgdata/geneseq/geneseq-emb/NA1986.DAT:
8: /SIDSI/gcgdata/geneseq/geneseq-emb/NA1987.DAT:
9: /SIDSI/gcgdata/geneseq/geneseq-emb/NA1988.DAT:
10: /SIDSI/gcgdata/geneseq/geneseq-emb/NA1989.DAT:
11: /SIDSI/gcgdata/geneseq/geneseq-emb/NA1990.DAT:
12: /SIDSI/gcgdata/geneseq/geneseq-emb/NA1991.DAT:
13: /SIDSI/gcgdata/geneseq/geneseq-emb/NA1992.DAT:
14: /SIDSI/gcgdata/geneseq/geneseq-emb/NA1993.DAT:
15: /SIDSI/gcgdata/geneseq/geneseq-emb/NA1994.DAT:
16: /SIDSI/gcgdata/geneseq/geneseq-emb/NA1995.DAT:
17: /SIDSI/gcgdata/geneseq/geneseq-emb/NA1996.DAT:
18: /SIDSI/gcgdata/geneseq/geneseq-emb/NA1997.DAT:
19: /SIDSI/gcgdata/geneseq/geneseq-emb/NA1998.DAT:
20: /SIDSI/gcgdata/geneseq/geneseq-emb/NA1999.DAT:
21: /SIDSI/gcgdata/geneseq/geneseq-emb/NA2000.DAT:
22: /SIDSI/gcgdata/geneseq/geneseq-emb/NA2001.DAT:
23: /SIDSI/gcgdata/geneseq/geneseq-emb/NA2002.DAT:
24: /SIDSI/gcgdata/geneseq/geneseq-emb/NA2003.DAT:
25: /SIDSI/gcgdata/geneseq/geneseq-emb/NA2004.DAT:

Prod. No. is the number of results produced by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4931	100.0	4931	ABL59333	Nucleotide sequence
2	1711.2	34.7	1186	ABL54414	Human PYRIN-8 cDNA
3	1616	32.8	3300	AAL47129	Pylin domain cDNA
4	826	16.8	2158	24 AB8787.9	Human cDNA encoding
5	492.8	10.0	506	21 AAS58757	RNA encoding novel
6	482.6	9.2	479	21 AAK76516	Human CRFX CRF2121
7	415	9.4	19199	22 AAK79995	Human immune/haema
8	410.4	8.3	26277	22 AAK70425	Human immune/haema

C 9	380.2	7.7	63606	24	ABL62910	Breast cancer rela
C 10	380.2	7.7	65608	24	ABL64414	Stomach cancer rei
C 11	380.2	7.7	65608	24	ABL67668	Oesophagus cancer
C 12	375.8	7.6	12312	25	AB274319	Secreted protein g
C 13	375.8	7.6	12312	25	AB274319	Human secreted pro
C 14	375.2	7.6	3857	22	AAC14323	Human PYRIN 1 (PYR
C 15	375.2	7.6	3857	25	ABX93556	Human cDNA encoding
C 16	375	7.6	16579	22	ABA20363	Human nervous syst
C 17	375	7.6	16579	22	AAL05755	Human reproductive
C 18	373	7.6	17792	22	AAS32727	Human genomic DNA
C 19	373	7.6	17792	22	AAS36099	Human cardiovascular
C 20	372.4	7.6	12245	22	AAK81897	Human immune/haema
C 21	371.8	7.5	7703	22	AAK90904	Human digestive sy
C 22	371.8	7.5	7703	22	AAK90945	Human digestive sy
C 23	371.8	7.5	7703	22	AAS11939	Human liver associ
C 24	371.8	7.5	7703	22	AAS11980	Human liver associ
C 25	371.8	7.5	7703	24	ABN96294	Human liver antige
C 26	371.8	7.5	7703	24	ABN96294	Human liver antige
C 27	367.4	7.5	32148	22	AAL64218	Human reproductive
C 28	364.2	7.4	23540	22	AAS28556	Genomic sequence #
C 29	364.2	7.4	23580	22	AAK97248	Human immune/haema
C 30	364.2	7.4	86080	24	ABQ88164	Human osteoblast d
C 31	364.2	7.4	96080	24	ABQ83561	Human cDNA differe
C 32	363.2	7.4	13673	22	ABA14719	Human nervous syst
C 33	359.6	7.3	32169	22	ABA19571	Human nervous syst
C 34	359.6	7.3	32169	22	ABL05917	Human reproductive
C 35	359.6	7.3	32169	23	ABL39448	Human testicular a
C 36	359.4	7.3	17123	22	AAK96738	Human immune/haema
C 37	358.4	7.3	11950	22	AAK93201	Human yesci Gene
C 38	356.4	7.2	32189	22	ABA19572	Human nervous syst
C 39	356.4	7.2	32189	22	AAL25916	Human reproductive
C 40	356.4	7.2	32189	23	ABL98480	Human testicular a
C 41	356.4	7.2	131971	22	AAK69257	Human immune/haema
C 42	355.6	7.2	121162	21	AAC66548	Human kinesin like
C 43	355.2	7.2	57248	24	ABK83563	Human cDNA differe
C 44	354.4	7.2	125791	22	AAS43104	Human Geströder re
C 45	353	7.2	114793	22	AAD08215	Human genome flu

ALIGNMENTS

RESULT 1

ABL59333
ID ABL59333 standard, cDNA, 4931 BP.

XX ABL59333:

XX 07-OCT-2002 (first entry)

XX Nucleotide sequence of human leucine-rich repeat protein HLRRL1.

XX Human; leucine-rich repeat; HLRRL1; proliferative disorder;
immune condition; atroposis; signal transduction; autoimmune disease;
haematopoietic cell disease; graft-versus-host disease; allergy; asthma;
cardiovascular disorder; neurological disease; pheromone;
pulmonary disease; chronic obstructive pulmonary disease;
allergic rhinitis; bronchial hyperresponsiveness; reproductive disease;
haematopoietic disease; platelet disorder; Bernard-Soulier syndrome;
inflammatory disorder; systemic lupus erythematosus;
cardiovascular disease; cancer; gene; ns.

XX Homo sapiens.

XX Key Location/Qualifiers

XX CDS 1320..2466

XX /tag a

XX /product "HLRRL1"

XX WO200252011-A2

XX 04 JUL 2002

XX

PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226685.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227209.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 03-SEP-2000; 2000US-0229509.
PR 06-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0232081.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 26-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.

PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249255.
PR 17-NOV-2000; 2000US-0249257.
PR 17-NOV-2000; 2000US-0249299.
PR 01-DEC-2000; 2000US-0251160.
PR 01-DEC-2000; 2000US-0251391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251949.
PR 11-DEC-2000; 2000US-0251990.
PR 05-JAN-2001; 2001US-0254097.
PR 05-JAN-2001; 2001US-0259678.
PR (HUYA-) HUMAN GENOME SCI :NC.

Rosen CA, Barash SC, Ruben SM;

WP1; 2001-483426/52.

Nucleic acids encoding human immune/haematopoietic antigen polypeptides,
useful for preventing, diagnosing and/or treating cancers and
metastasis -

Disclosure: SEQ ID NO 25807; 307.ipp + Sequence Listing; English.

AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (i)
amino acid sequences given in AAK62170 to AAK91921. (i) have catalytic
activity, and can be used in gene therapy and vaccine production. (i)
proteins and polynucleotides may be used in the prevention, diagnosis and
treatment of diseases associated with inappropriate (i) expression. For
example, they may be used to treat disorders associated with decreased
expression by rectifying mutations or deletions in a patient's genome
that affect the activity of (i) by expressing inactive proteins or to
supplement the patient's own production of (i). Additionally, (i)
polynucleotides may be used to produce the secreted (i), by inserting
the nucleic acids into a host cell and culturing the cell to express the
protein. (i) proteins and polynucleotides may be used to prevent,
diagnose and treat immune/haematopoietic-related diseases, especially

CY	4384	----- AAACAAATAAATAATCAACACGGCGTGGCGCGTG	4828	PR	14-AUG-2000	2000US-0205266
EB	13236	ANTTCATAGGTTTCAGTGAAATAAATAAAGAGATCTCAGGTTTGGCAGAGTG	13155	PR	14-AUG-2000	2000US-0225267
QY	4829	GCTGATGCTGTAAACCGACACTTTTGGAGGCTGAGGCTGGTGATCAAC	4866	PR	14-AUG-2000	2000US-0225268
DE	13156	TCTCATCTTTGTAATCTCCAGCAATTTAGGAGCTTAGGAGGATGCTTGAGCCAG	13215	PR	14-AUG-2000	2000US-0225270
QY	4837	GAGATGAGACATCTCTGGGAC TGGTAAAGTCTGTCTACTAAATAATAC-AATAA	4744	PR	14-AUG-2000	2000US-0225447
DE	13216	AAGTTTGACATCAGCTTAGGCAATCGTGAAGA TATATATACAAAAGTACAAAATA	13275	PR	14-AUG-2000	2000US-0225477
QY	4845	TGACTTGGGCTGTGGTGTCACCTGTATATAGAGTACTACAGAGTTGATGCAAG	4804	PR	14-AUG-2000	2000US-0225478
DE	13236	TGACTTGGGCTGTGGTGTCACCTGTATATAGAGTACTACAGAGTTGATGCAAG	13335	PR	14-AUG-2000	2000US-0225479
QY	4855	AATACATTCACACAGGAGGAGGAGGTTTACAT TACGCGAATGATGCGACTGATCCCA	4864	PR	14-AUG-2000	2000US-0225481
DE	13336	GATCATTACGCTGGGAGGTCAGTCTGTAAATTAATGATGATTCCTCCTACTACATCTA	13395	PR	14-AUG-2000	2000US-0225482
QY	4865	GCTTGGGCGACAGAGTGGATTCCTATCTGAAAAAATAAATAAATAAATAAATAA	4924	PR	14-AUG-2000	2000US-0225483
DE	13396	GCTTGGGCGACAGAGTGGATTCCTATCTGAAAAAATAAATAAATAAATAAATAA	13455	PR	14-AUG-2000	2000US-0225484
QY	4925	AAAAA 4929		PR	14-AUG-2000	2000US-0225485
DE	13456	AAAAA 13460		PR	14-AUG-2000	2000US-0225486
RESULT 4						
AAAT-42872						
CO	AAAT-42872 standard; DNA; 26277 BP.					
XX	AAAT-42872;					
XX	06 NOV-2001 (first entry)					
XX	Human immune/haematopoietic antigen sequence SEQ ID NO:24240.					
XX	Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;					
XX	Cytostatic; gene therapy; vaccine; not identified.					
CS	Homo sapiens.					
XX	W020017182-A2.					
XX	09-AUG-2001.					
XX	13 JAN-2001; 2001WC-US01354.					
XX	31 JAN-2000; 2000US-0179065.					
XX	04 FEB-2000; 2000US-0180628.					
XX	24 FEB-2000; 2000US-0184664.					
XX	02 MAR-2000; 2000US-0186353.					
XX	14 MAR-2000; 2000US-0188874.					
XX	17 MAR-2000; 2000US-0190376.					
XX	18 APR-2000; 2000US-0198123.					
XX	19 MAY-2000; 2000US-0205515.					
XX	07 JUN-2000; 2000US-0209467.					
XX	28 JUN-2000; 2000US-0214866.					
XX	30 JUN-2000; 2000US-0215135.					
XX	07 JUL-2000; 2000US-0216647.					
XX	07 JUL-2000; 2000US-0216880.					
XX	11 JUL-2000; 2000US-0217487.					
XX	11 JUL-2000; 2000US-0217496.					
XX	14 JUL-2000; 2000US-0218293.					
XX	26 JUL-2000; 2000US-0220963.					
XX	26 JUL-2000; 2000US-0220964.					
XX	14 AUG-2000; 2000US-0224518.					
XX	14 AUG-2000; 2000US-0224519.					
XX	14 AUG-2000; 2000US-0225213.					
XX	14-AUG-2000; 2000US-0225214.					

XX AB274319;
 AC
 AT 12-MAY-2003 (first entry)
 CE Secreted protein gene 236 genomic fragment HNTNH73, SEQ ID NO:1456.
 XX Human; secreted protein; cancer; tumor; hyperproliferative disorder;
 KW autoimmune disorder; inflammation; angiodysplastic diseases; AIDS;
 KW acquired immunodeficiency syndrome; hepatitis; anaemia; wound healing;
 KW drug screening; chromosome identification; chromosome mapping;
 KW cytostatic; gene therapy; anti-inflammatory; immunomodulator; anti-HIV;
 KW antianemic; vulvectomy; gene; ds.
 XX Homo sapiens.
 CS
 XX W02000273013-A2.
 XX
 XX 26-MAR-2002; 2002MC-US09372.
 XX
 XX 22-MAR-2001; 2001US-278652P.
 PP 12-SEP-2001; 2001US-0953082.
 PR 12-SEP-2001; 2001US-0953083.
 XX
 XX (HUMAN) HUMAN GENOME SCI. INC.
 XX Rosen CA, Ruben SM;
 XX WF1; 2003-040578703.
 XX New human secreted proteins and nucleic acids, useful for detecting or
 PP treating cancer or other hyperproliferative disorders, autoimmune
 PP disorders, inflammatory disorders, HIV disease, hepatitis or anemia.
 XX Disclosure, Page 2059-2062; 2474pp, English.
 XX
 XX AR273281-AR273697 represent cDNAs corresponding to 301 human secreted
 CC protein genes, and ABP00947-ABP1163 represent the proteins they encode.
 CC AR273698-AR274687 represent human secreted protein genomic fragments. The
 CC invention also encompasses antibodies specific for the secreted proteins.
 CC the use of the secreted proteins in drug screening and recombinant
 CC vectors and host cells comprising a nucleic acid of the invention. The
 CC secreted proteins are thought to be involved in biological activities
 CC associated with cellular signalling, cellular differentiation, cell
 CC migration, proinflammatory activation and neuroendocrine activity. The
 CC secreted proteins, nucleic acids encoding them, antibodies or antibody
 CC fragments specific for the secreted proteins, and modulators of protein
 CC activity are useful for diagnosing or treating cancers or other
 CC hyperproliferative disorders. Additionally, the secreted proteins and
 CC their nucleic acids may also be used in the treatment of autoimmune
 CC disorders, inflammatory disorders, diseases involving angiogenesis, AIDS
 CC (acquired immunodeficiency syndrome), hepatitis, anaemia, and to promote
 CC wound healing. Nucleic acids of the invention may be used for chromosome
 CC identification, chromosome mapping, in gene therapy, for identifying
 CC individuals from minute biological samples, as hybridisation probes, and
 CC as molecular weight markers. The present sequence represents a human
 CC secreted protein genomic fragment referred to in the disclosure of the
 CC invention.
 XX
 SQ Sequence 12212 BP: 3311 A: 2770 C: 2065 G: 1166 T: 5 other;
 Query Match: 7.6%; Score 375.4; DB 25; Length 12212;
 Best Local Similarity 60.7%; Pred. No. 7 in 64;
 Matches 846; Conservative 0; Mismatches 492; Indels 55; Gaps 12;
 3569 TTTTATTTTGTAGTACAGACAGGTTTATATGTTAG-CAGSAGTGTCTGGAT 3687
 DB 375 TTTCTCTGCTCCAGCTCCCGAGTAGTGTGGATATAGSAGAGACACACCTCGGC 434
 3629 TAACTTTTGTATTTTGTAGTACAGACAGGTTTATATGTTAG-CAGSAGTGTCTGGAT 3687

DB 435 TAA TTTTGTATTTTGTAGTACAGACAGGTTTATATGTTAGTGTGTGCTGTCTGAAA 492
 QY 3688 CTATGAGACTCA TGAATGCTGTGCTTTGGCTTCCCAAAAGTGTGGATTAAGAGCTTG 3745
 DB 493 CCGCTGAGCTTGAATGATGACCGGCTAACCTCCCAAGTGTAGGATTACAGGATG 562
 QY 3746 ACCCACTGTATGCACTTACATCAAGCTTTT ----- 3777
 DB 563 AGTACAGGATATGACCAATTTATTTTATTGTGTAGAGTTTACCTGAAATAAAAAA 612
 QY 3778 --AAAGGATTTTGTGCTGATGCTGTGCTGACACTTAACTCAGCACTTTTGGGA 3835
 DB 613 AAAAGAGATTATGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 672
 QY 3836 GCGTGAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3895
 DB 673 GCGTGAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 732
 QY 3896 AAAGCTGCTATGCTTAAATATAAAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3955
 DB 733 AA--AGGCTATGCTTAAATATAAAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 790
 QY 3956 TCGAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4015
 DB 791 TCGAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 850
 QY 4016 TCGAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4075
 DB 851 TCGAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 910
 QY 4076 AAATAAATAAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4135
 DB 911 AAATAAATAAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 970
 QY 4136 CTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4195
 DB 971 AAGAGCAAGTTCGAGTGTGAGCTTTTCTGAGGAGGCTGCTGCTGCTGCTGCTGCTGCTG 1030
 QY 4196 AAATAAATAAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4255
 DB 1031 CATTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1090
 QY 4256 GTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4308
 DB 1091 CAGCTCAAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1150
 QY 4309 CTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4367
 DB 1151 TACAGCAAGTTCGAGTGTGAGCTTTTCTGAGGAGGCTGCTGCTGCTGCTGCTGCTGCTG 1212
 QY 4368 TTTTGTAGATTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4427
 DB 1211 CTTTGTAGATTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1270
 QY 4428 AAATAAATAAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4487
 DB 1271 CATTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1330
 QY 4488 GAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4542
 DB 1331 GAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1390
 QY 4543 GAAATAAATAAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4602
 DB 1391 CTTTGTAGATTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1450
 QY 4602 AAATACAGACAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4661
 DB 1451 AAAGCTGCTATGCTTAAATATAAAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1510
 QY 4662 TGAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4718
 DB 1511 GAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1570

QY 1494 TGTGGAGGAGAACGGCCACGGAGCTGTTTCTTACAGCTTATTGGAGAGAGCTG 1553
 DE 1093 GACTGGGAGAGGGAGGGGGAGAGACATTTCTGTAGCAGGCTTCATCAGAAAGAGCTG 1152
 QY 1554 CTGCTGAGCTATCTTGTCTCATACACAGGAGGAGGCTTGTGTGAGAGAGCTGACCGT 1613
 DB 1153 CTGCGGAGGCTTCTGCTCATACACAGGAGGAGGCTTGTGTGAGAGAGCTGACCGT 1613
 QY 1614 CTGCTGAGAGCCAGGAGGAGCTGAGAGCTGAGAGCTTGTGTGAGAGAGCTGACCGT 1613
 DB 1213 TGTGTGAGAGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1272
 QY 1674 TACTTCTACAGTATTTCCAAATTCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1733
 DB 1273 TACTTCTACAGTATTTCTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1332
 QY 1734 GAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1793
 DB 1333 GAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1392
 QY 1794 TGCCTCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1853
 DB 1393 GAGTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1452
 QY 1854 GAGTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1913
 DB 1453 GAGTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1512
 QY 1914 GAGTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1973
 DB 1513 GAGTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1572
 QY 1974 GAGTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2033
 DB 1573 GAGTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1632
 QY 2034 GAGTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2097
 DB 1633 GAGTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1696
 RESULT 15
 ID AEX93556 standard; cDNA; 3857 BP.
 AC AEX93556;
 CT 22-MAY-2003 (first entry)
 CE Huma cDNA encoding PYRIN 1.
 KW Human; ss: Gene: nucleoside binding site; pyrin domain; NBS-1; PYRIN-1;
 KW caspase 1; anti-inflammatory; apoptosis; ASC; NF- κ B; nuclear factor κ B;
 KW LRR; leucine rich repeat; inflammatory disorder; familial cold urticaria;
 KW arthritis; inflammatory bowel disease; Crohn's disease; asthma; allergy;
 KW ulcerative colitis; rheumatoid arthritis; Lyme disease; Grave's disease;
 KW insulin-dependent diabetes; multiple sclerosis; contact dermatitis;
 KW psoriasis; graft versus host disease; food allergy; conjunctivitis;
 KW chronic obstructive pulmonary disease; food allergy;
 KW helminthic infection; leishmaniasis; viral infection; HIV infection;
 KW bacterial infection; tuberculosis; leprosy; infectious lq44.
 XX Homo sapiens.
 XX Key Location/Qualifiers
 FT 139..3243
 CD /tag= a
 FT /product= "PYRIN-1"
 XX GS2002187922-A1.
 PN 12-DEC-2002.

XX 22-APR-2002; 2002US 0137516.
 PR 17-FEB-2000; 2000US 0556067.
 PR 01-SEP-2000; 2000US 0653901.
 PR 26-SEP-2001; 2001US 0949951.
 PR 20-DEC-2001; 2001US 0027629.
 PA (BERT/) BERTIN J.
 PA (MANU/) MANU: G A.
 XX Bertin J, Manu J GA.
 XX WP1: 2003 328763/11.
 DR P-PSDB; ABL04503.
 XX Identification of compound that binds to polypeptide (for e.g. a PYRIN protein), useful for treating disorders associated with inappropriate apoptosis, for e.g. inflammatory disorder.
 XX Disclosure; Fig 4: 74pp; English.
 PS The invention relates to a new method for the identification of a compound that binds to a human PYRIN-1 (an apoptotic signaling molecule appearing as ABL04503) comprising contacting the polypeptide or a cell expressing the polypeptide to a test compound and determining whether PYRIN-1 binds to the test compound, or determining the effect of the test compound on the activity of PYRIN-1 in a similar manner.
 CC compounds are isolated which modulate the binding of PYRIN-1 to ASC comprising testing the compounds against the binding of PYRIN-1 to ASC both PYRIN-1 and ASC (not defined), modulate the activity of NF- κ B (nuclear factor κ B), and which modulate the ASC-mediated activation of NF- κ B (comprising: (a) measuring the binding of a test compound to the LRR (leucine rich repeat) domain of PYRIN-1; and (b) measuring the activation of NF- κ B in a cell expressing ASC and PYRIN-1 in the presence and absence of the LRR domain binding compound). The candidate modulators may be useful for treating an inflammatory disorder.
 CC Also included is a method (ME) for identifying a modulator of caspase-1 activity. The identified compound can be used to modulate the polypeptide's activity (ASC and NF- κ B activities in a patient). By modulating the expression or activity of the polypeptide (PYRIN-1) and ASC, a disorder associated with inappropriate apoptosis (inflammatory disorders e.g. familial cold urticaria, arthritis, inflammatory bowel disease, Crohn's disease, ulcerative colitis, rheumatoid arthritis, Lyme disease, insulin-dependent diabetes, multiple sclerosis, Grave's disease, contact dermatitis, psoriasis, graft versus host disease, asthma, chronic obstructive pulmonary disease, allergies (e.g. food allergy), conjunctivitis, helminthic infection (e.g. leishmaniasis), viral infections such as HIV infection, and bacterial infections such as tuberculosis and leprosy) can be treated. Human NBS-1 protein (nucleoside binding site protein 1) contains a pyrin domain which was used to identify genomic sequences containing the PYRIN-1 gene. The gene for PYRIN-1 is located on chromosome 1q44. The present sequence encodes human PYRIN-1.
 XX Sequence 3857 BP; 958 A; 931 C; 1023 G; 945 T; C other;
 QY Query Match 7.6%; Score 375.2; DB 25; Length 3857;
 DB Best Local Similarity 60.2%; Pred. No. 7.7e-64;
 DB Matches 713; Conservative C; Mismatches 378; Indels 97; Gaps 2;
 QY 985 TGGCTCTGGAGTCTTCTGTCTACTCCAGAAAGATCCCGAGGAACCTACAGGAC 1044
 DB 499 TACCTTCGAGATCTCTCTATTGTGAAATGAAGAGAGATTACCGTAGAGATACAGAA 558
 QY 1045 TATGTCCGAGGAAATTCGGGCTCATGAGACCGCAATCGCCCTAGGGAAATGTGTC 1104
 DB 559 TACGTGAGAGAGATTCAGTTCATGATGAGACAGCAATCCCGTCTGGTGAAGTGTG 618
 QY 1105 AATCTCAGGACCGGTACACCGGCTCTGTGTGTGAGGAGGACTCAAAACCCATCGAG 1164
 DB 619 AGCCTCAAGAAACCGGTACACAGGACTGCTCTCATCAAGGAGGACCGGAGCCAGGAG 678

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

GM nucleic - nucleic search, using sw model

Run on: October 30, 2003, 21:38:28, Session time 17529 Seconds
Without alignments: 11508.093 Million cell updates/sec

Title: US-10-028-374-1

Perfect score: 4931

Sequence: 1 ccacgcgcgcgaacatgggc.....dadaaaadaaaadaaaa 4931

Scoring table: IDENTITY NUC

Gap: 10.0, Gapext: 1.0

Searched: 2893711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 577422

Minimum DB seq length: 9

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database:

GenBank:

1: gb_ba:

2: gb_hg:

3: gb_in:

4: gb_ov:

5: gb_ov:

6: gb_ov:

7: gb_ph:

8: gb_ph:

9: gb_ph:

10: gb_ph:

11: gb_ph:

12: gb_ph:

13: gb_ph:

14: gb_ph:

15: gb_ph:

16: gb_ph:

17: gb_ph:

18: gb_ph:

19: gb_ph:

20: gb_ph:

21: gb_ph:

22: gb_ph:

23: gb_ph:

24: gb_ph:

25: gb_ph:

26: gb_ph:

27: gb_ph:

28: gb_ph:

29: gb_ph:

30: gb_ph:

31: gb_ph:

32: gb_ph:

33: gb_ph:

34: gb_ph:

35: gb_ph:

36: gb_ph:

37: gb_ph:

38: gb_ph:

39: gb_ph:

40: gb_ph:

41: gb_ph:

42: gb_ph:

43: gb_ph:

44: gb_ph:

45: gb_ph:

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	4931	100.0	4931	6	AX684291 Sequence
2	1733.2	34.7	3827	9	AY095146 Homo sapi
3	1733.2	34.7	3827	9	AY154467 Homo sapi
4	1733.2	34.7	3827	9	AY154467 Homo sapi
5	1710	34.7	3827	9	BC028069 Homo sapi
6	1697	34.4	3563	9	AY116205 Homo sapi
7	1697	34.4	3563	9	AY116205 Homo sapi
8	1616	32.8	3128	6	AX417214 Sequence
9	1616	32.8	3128	6	AX459869 Sequence
10	1577.6	32.0	3195	9	AY116206 Homo sapi
11	1524.8	30.9	4221	9	AY116207 Homo sapi
12	1226.2	24.9	14730	9	AC008753 Homo sapi
13	826	16.8	2158	6	AX575503 Sequence
14	783.6	15.9	2494	9	AF231021 Homo sapi
15	723.8	14.5	22344	2	AC104193 Rattus no
16	662.6	13.4	21827	2	AC079499 Mus muscu
17	507.4	10.3	98758	9	HS223H9
18	460.8	9.3	19879	9	AC112215 Homo sapi
19	457.6	9.3	186734	9	AC104446 Homo sapi
20	457.6	9.3	215929	9	AC105934 Homo sapi
21	454.6	9.2	13592	9	AC105934 Homo sapi
22	452.8	9.2	59836	9	AC105934 Homo sapi
23	452.8	9.2	180283	9	AC105934 Homo sapi
24	451.8	9.2	20724	9	AC105934 Homo sapi
25	446.2	9.0	42265	9	AC105934 Homo sapi
26	446.2	9.0	179934	9	AC105934 Homo sapi
27	445.4	9.0	92766	2	AF322454 Homo sapi
28	445.4	9.0	101685	9	HS019622 Homo sapi
29	435	8.8	195952	9	AC12184 Homo sapi
30	430.2	8.7	12132	9	AC006441 Homo sapi
31	429.6	8.7	165617	9	AF165926 Homo sapi
32	429.6	8.7	177782	9	AC117532 Homo sapi
33	429.4	8.7	116762	9	AC108004 Homo sapi
34	428	8.7	107717	9	AC137055 Homo sapi
35	428	8.7	108139	2	AC004285 Homo sapi
36	427.6	8.7	100000	9	AF000531 Homo sapi
37	426	8.6	10127	9	AC104134 Homo sapi
38	426	8.6	160335	9	AF000729 Homo sapi
39	426	8.6	177592	9	AC093814 Homo sapi
40	426	8.6	162769	9	AC093814 Homo sapi
41	426	8.6	189406	9	AC112717 Homo sapi
42	425.2	8.6	108921	2	HS112K5 Homo sapi
43	422.2	8.6	133865	2	AC137962 Homo sapi
44	421	8.5	137786	9	AF450326 Homo sapi
45	420.4	8.5	196686	9	AC010328 Homo sapi

ALIGNMENTS

RESULT:

AX684291

LOCUS

DEFINITION

Sequence 1 from Patent WO2005011

AX684291

ACCESSION

AX684291.1

KEYWORDS

SOURCE

ORGANISM

Homo sapiens (human)

REFERENCE

AUTHORS

TITLE

Human leucine rich repeat containing protein, Lrrbml, expressed

predominately in bone marrow

PAT 29-MAR-2003

Pred. No. is the number of results predicted by chance to have a

QY 1801 AGCAGCAGCTGGAGGGTGGGGGCTGTCAGACAGACGTCAGGACCACTGCAGTGT 1860
DB 1801 AGCAGCAGCTGGAGGGTGGGGGCTGTCAGACAGACGTCAGGACCACTGCAGTGT 1860
QY 1861 ACATGCTCTACCTGCTGAGTCTGATCAACGCAAGCTGGGGGCTGGGCTCCAGCCCC 1920
DB 1861 ACATGCTCTACCTGCTGAGTCTGATCAACGCAAGCTGGGGGCTGGGCTCCAGCCCC 1920
QY 1921 CACCCACACAGAGAGGTTGTGCTCTGGCGGACAGATGGGCTCTGGAAATCGAATAATCC 1980
DB 1921 CACCCACACAGAGAGGTTGTGCTCTGGCGGACAGATGGGCTCTGGAAATCGAATAATCC 1980
QY 1981 TATTTCAGGAGCAGGACCTCCGAAACAGCGCTAGACGCGAAAGACGCTCTGCTTCC 2040
DB 1981 TATTTCAGGAGCAGGACCTCCGAAACAGCGCTAGACGCGAAAGACGCTCTGCTTCC 2040
QY 2041 TCACATGAACATCTTCCAGAGGATTAACATGTGAGAGAGGCTTCTGGCATTCACCA 2100
DB 2041 TCACATGAACATCTTCCAGAGGATTAACATGTGAGAGAGGCTTCTGGCATTCACCA 2100
QY 2101 GCGGCTTCTGTTGACCTCTGAGAGAGACGAGGATGACATGGAGAGAGTCTCT 2160
DB 2101 GCGGCTTCTGTTGACCTCTGAGAGAGACGAGGATGACATGGAGAGAGTCTCT 2160
QY 2161 GCTGGAGGTTCTCCCGACATCAAGATGAGCTGTGGAGTGGATCCAAAGCAAGCTC 2220
DB 2161 GCTGGAGGTTCTCCCGACATCAAGATGAGCTGTGGAGTGGATCCAAAGCAAGCTC 2220
QY 2221 AGACCGACGCTCGACCTCGACAGAGGCTCTGAGTGTGAGAGAGGCTTGTACAGAGA 2280
DB 2221 AGACCGACGCTCGACCTCGACAGAGGCTCTGAGTGTGAGAGAGGCTTGTACAGAGA 2280
QY 2281 TCACAGAGGAGGAGTATTATCCAGACGCGCTGAGTACCTCCAGGTGATCGTGGTACGA 2340
DB 2281 TCACAGAGGAGGAGTATTATCCAGACGCGCTGAGTACCTCCAGGTGATCGTGGTACGA 2340
QY 2341 AATTCGCTCCAGATGGAGACATGGTCTGCTGCTGCTGCTGCTGAGAGGCTGAGAGCG 2400
DB 2341 AATTCGCTCCAGATGGAGACATGGTCTGCTGCTGCTGCTGCTGAGAGGCTGAGAGCG 2400
QY 2401 CCGAGTGTGCTACTGTATGGCGCAGCTACAGTGGGAGCGGAGACCGGGTGAAGGT 2460
DB 2401 CCGAGTGTGCTACTGTATGGCGCAGCTACAGTGGGAGCGGAGACCGGGTGAAGGT 2460
QY 2461 GCTCCGAGAGCGCACACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2520
DB 2461 GCTCCGAGAGCGCACACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2520
QY 2521 CTTACAGTGAACATCTGGCAGCGGCTGTGTGCAAGATCGAAGCTGATAGAGCTGTCTC 2580
DB 2521 CTTACAGTGAACATCTGGCAGCGGCTGTGTGCAAGATCGAAGCTGATAGAGCTGTCTC 2580
QY 2581 TGTACCGAAATCGCTGGGAGCGGGGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2640
DB 2581 TGTACCGAAATCGCTGGGAGCGGGGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2640
QY 2641 CCAACTGCAAACTTCAGAACTGAGTAAATTTATCATATATAAATGATATTTTGA 2700
DB 2641 CCAACTGCAAACTTCAGAACTGAGTAAATTTATCATATATAAATGATATTTTGA 2700
QY 2701 TAAATATATGGCCAGGTATGATGGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2760
DB 2701 TAAATATATGGCCAGGTATGATGGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2760
QY 2761 GATGGGAGGATCATCTTACCCAGGAGTTCAAGACAGGCTGGCTAAATGATGAAACCC 2820
DB 2761 GATGGGAGGATCATCTTACCCAGGAGTTCAAGACAGGCTGGCTAAATGATGAAACCC 2820
QY 2821 CATCTCTACTAAATAACCAAAATGAGCGAGGATGCTGAGAGAGGCTGCTGAGAGAGG 2880
DB 2821 CATCTCTACTAAATAACCAAAATGAGCGAGGATGCTGAGAGAGGCTGCTGAGAGAGG 2880

QY 2881 TACTCAGGAGCGCAAGGAGGAGTTCCTCAACCCAGGAGGACAGGTTGTGGCTGAA 2940
DB 2881 TACTCAGGAGCGCAAGGAGGAGTTCCTCAACCCAGGAGGACAGGTTGTGGCTGAA 2940
QY 2941 GAGTGCCGCATCTCCAGCTCAGCTCGAGGACCTCTCTCAGCTCTCATAGCAATAA 3000
DB 2941 GAGTGCCGCATCTCCAGCTCAGCTCGAGGACCTCTCTCAGCTCTCATAGCAATAA 3000
QY 3001 GAATTTGCAAGAGTGGATCTCAGTGGCAAGGGCTTGATTCAGGACCATGATGCTGCT 3060
DB 3001 GAATTTGCAAGAGTGGATCTCAGTGGCAAGGGCTTGATTCAGGACCATGATGCTGCT 3060
QY 3061 TTCCGAGGCGCTTGGCGCATCCCAATGCGAGGCTGAGATGATTCAGTTGAGCAAGTGTCA 3120
DB 3061 TTCCGAGGCGCTTGGCGCATCCCAATGCGAGGCTGAGATGATTCAGTTGAGCAAGTGTCA 3120
QY 3121 GCTTGAATTCGCGGCTTGTTCAGAGATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3180
DB 3121 GCTTGAATTCGCGGCTTGTTCAGAGATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3180
QY 3181 TGAGTTGGAGCTTGACAGAGAAATGCACTGGAGGATTTGGGCTGAGGTTACTATGCAAGG 3240
DB 3181 TGAGTTGGAGCTTGACAGAGAAATGCACTGGAGGATTTGGGCTGAGGTTACTATGCAAGG 3240
QY 3241 ACTGAGGCACTGAGCTTTCAGACTACGACTTTGTGGCTGAAAGTCTCGGCTCTCACTGC 3300
DB 3241 ACTGAGGCACTGAGCTTTCAGACTACGACTTTGTGGCTGAAAGTCTCGGCTCTCACTGC 3300
QY 3301 TGGTGGCTGTGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3360
DB 3301 TGGTGGCTGTGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3360
QY 3361 CCGTGAAGCTGAAATGAGCTTGGGAGCTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3420
DB 3361 CCGTGAAGCTGAAATGAGCTTGGGAGCTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3420
QY 3421 TCCGCTGCTGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3480
DB 3421 TCCGCTGCTGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3480
QY 3481 CCAATCCATGACGAGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3540
DB 3481 CCAATCCATGACGAGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3540
QY 3541 AATCTTAGCTACTATCATCAGCCCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 3600
DB 3541 AATCTTAGCTACTATCATCAGCCCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 3600
QY 3601 ATTACAGGCGCGGCGCACAGGCTGAGTAACTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 3660
DB 3601 ATTACAGGCGCGGCGCACAGGCTGAGTAACTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 3660
QY 3661 TCATCATGTTAGCAGGATGCTCTGATCTACTGACTCATGATCTGCTGCTGCTGCTGCTGCTGCT 3720
DB 3661 TCATCATGTTAGCAGGATGCTCTGATCTACTGACTCATGATCTGCTGCTGCTGCTGCTGCTGCT 3720
QY 3721 CCAAGTGTGAGATTTACAGGCTTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3780
DB 3721 CCAAGTGTGAGATTTACAGGCTTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3780
QY 3781 AGGATTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3840
DB 3781 AGGATTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3840
QY 3841 AGGTGGAGGATCATCTGAGGCTGGGAAATTCAGAGCAGGCTGACCAACATGTAAGAAACA 3900
DB 3841 AGGTGGAGGATCATCTGAGGCTGGGAAATTCAGAGCAGGCTGACCAACATGTAAGAAACA 3900
QY 3901 CCGCATCTGCTACTAAATAACAAAGTACCCAGGATGCTGGTGCATGCTTATATATGCA 3960
DB 3901 CCGCATCTGCTACTAAATAACAAAGTACCCAGGATGCTGGTGCATGCTTATATATGCA 3960
QY 3961 GCTACTGCGAGGCTGAGGAGAGAAATCCGTTGAAACCCGGGAGGCTGGAAGTTGCTGTGA 4020


```

QY 3111 GGAGTGTTCAGCTGGAGTCCGGGCTTTTCAGCGATAGCTTCTGTCTGTCGACCAACC 3170
EI ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2596 GGAAGTGTTCAGCTGGAGTCCGGGCTTTTCAGCGATAGCTTCTGTCTGTCGACCAACC 2655
QY 3171 CACATCTGTTGAGTGGACCTGACACAAATGCACTGAGAGNTTGGGCTGAGGTTAC 3230
EI ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2656 CACATCTGTTGAGTGGACCTGACACAAATGCACTGAGAGNTTGGGCTGAGGTTAC 2715
QY 3231 TATGCGAGGACTGAGGACCACTGCTGAGACTACGACTTGTGCTGCTGAGGATCTGCC 3290
EI ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2716 TATGCGAGGACTGAGGACCACTGCTGAGACTACGACTTGTGCTGCTGAGGATCTGCC 2775
QY 3291 GCTCACTCTGCTGCTGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3350
EI ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2796 GCTCACTCTGCTGCTGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2855
QY 3351 GAGAGCTGAGCTGAGCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAG 3410
EI ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2836 GAGAGCTGAGCTGAGCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAG 2895
QY 3411 GCTCAGGACTCCAGCTCCAGCTCCAGCTCCAGCTCCAGCTCCAGCTCCAGCTCC 3454
EI ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2836 GCTCAGGACTCCAGCTCCAGCTCCAGCTCCAGCTCCAGCTCCAGCTCCAGCTCC 2919

FEATURES             Source
AY154467              3827 bp     mRNA     linear     PRI: 20 FEB 2003
LOCUS                 Homo sapiens NALP12 (NALP12) mRNA, complete cds.
ACCESSION              AY154467
VERSION                AY154467.1 GI:28416377
KEYWORDS               Homo sapiens (human);
SOURCE                 Homo sapiens
ORGANISM               Bakaliyca; Metazoa; Chordata; Clariata; Vertebrata; Euteleostomi;
                       Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE              1 (bases 1 to 3827)
AUTHORS                Tschoep, J., Martinon, F. and Burns, K.
TITLE                 NALP12: a novel protein family involved in inflammation
JOURNAL                Nat. Rev. Mol. Cell Biol. 4 (2003) 91-104 (2003)
MEDLINE                22451042
PubMed                12563287
REFERENCE              2 (bases 1 to 3827)
AUTHORS                Martinon, F., Hofmann, K. and Tschoep, J.
TITLE                 Direct Submission
JOURNAL                Submitted (25-SEP-2002) Institute of Biochemistry, University of
                       Lausanne, ch des Boveresses 155, Epalinges, VD 1066, Switzerland
                       location/Qualifiers
                       1..3827
                       /organism="Homo sapiens"
                       /mol_type="mRNA"
                       /db_xref="taxon:9606"
                       1..3827
                       /gene="NALP12"
                       177..3362
                       /gene="NALP12"
                       /notes="member of the NALP protein family involved in
                       inflammation; contains NACHT, LRR and PYD domains"
                       /codon_start=1
                       /product="NALP12"
                       /protein_id="AAC18163"
                       /db_xref="GI:28416377"
                       /translations="MURTAGGGLGALSYLELDAVLEIKPKLYLGTATGEGGKIP
                       WGMKAGLENAQLLTHFGEAMRLALSTFERINRKLWEROEDLVETPPGG
                       PSSGNGSTCLLEVSVTFRKQFQTYPDVVMKRNINEDNRARIGECVNLSHRYTRL
                       LLVKEHNPQVQOCLLDGTGSHATVGHQASHIKLTLFEPUREEPPTVMQGA
                       AGTGLKMLAKWMLWDADKLPGGREDILFYINRKNYGSATGECMQBLIPSCPEPS
                       APQELIRPFRPLFLDGGDELKPSFHPQSPWELCWHPKPTSLINLIRKRLCP
                       ELSLITTPALEK:HRLEHPRVEIFGSEPRERYFYRFFNAEQAGVKNYVR
                       DNEPTMTGVLVPCWVCTLQQQLEGGRLRQTSPTTAVVM:YVLSLQCPKPGAP
                       RLQPPNQRGLGSAAGLNCKLFEHQDRKGLQEDVSFA:KVNIFQKINCEIR
                       VYSFHLSTGEFFAAVYVILDEGRGAGPQQRVTELLTFYAFSEBSFIALTSRPLFLG

```

```

INNEETRSDEKSLCKVSH:KYDILWIOGSKAQSDGSTLQGGLEFFSLYEIQEEF
FLOALSHFOVIVSN:ASKVEHVMVSSFLKCRSAVLHYLYGATGAGGDRARCSA
GARTYVLPBRTVLLDVAEHLAACTNPNELELSYPAVAGSRGVKXGCOLPBP
NKKLNIRKRCRISSSACECLSAALANKNTERMDLSGNGVGFEGMMLRGEHRHC
QFQMTQKKQCESGACXVASVGTNPHJVELDUTGNALDELJLRHJCOGLRHPVC
PRTIMKICRLTAACCELASTSVNQSRELDLSNELGDLVLLCE:PIPTCK
IQTLRLG:PLRSSAACETSVVQAHNRLELDLSFNDLQWKLIAELHLPACPR
QKALDSQNTAKAVENYFTLGINCTUDLYLTNNALGDTGVRLLKPLSHPOCKLR
VWIKFENCLNKVTHSELAALSVTKTYLDIG"
BASE COUNT      454 a 1245 c 1122 g 246 t
ORIGIN
Query Match      34.7% Score 113.27 DB 9; Length 3827;
Best local similarity 61.3% Pval Exp. 0;
Matches 2241; Conservative 3; Mismatches 458; Gaps 3;
QY 940 GATACCGATTTATGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 999
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
465 GATACCGATTTATGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 524
QY 1000 GCTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1059
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
525 GCTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 584
QY 1060 GTCGCGTATGAGAGACGCAATGTCGCTGAGGGATGTGTCAATCTGAGCAACGG 1119
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
585 TTCCGGTTCATGGAAACCGCAATGTCGCTGAGGGATGTGTCAATCTGAGCAACGG 644
QY 1120 TACAGCGCTTCTGTGTGTGAGAGAGTCTCAACACCCATGCGAGTCCAGCAGCGTT 1179
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
645 TACAGCGCTTCTGTGTGTGAGAGAGTCTCAACACCCATGCGAGTCCAGCAGCGTT 704
QY 1180 CTGACACAGCGCGGTJACACCGAGGAGCGCTGCGACACCCAGGCTAGCCCATCAAGATA 1239
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
705 CTGACACAGCGCGGTJACACCGAGGAGCGCTGCGACACCCAGGCTAGCCCATCAAGATA 764
QY 1240 GAGATCTCTTTTAGGATATACATAGAGAGGATGCTGAGGACGCGCGACCGTGTGTGCA 1299
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
765 GAGATCTCTTTTAGGATATACATAGAGAGGATGCTGAGGACGCGCGACCGTGTGTGCA 824
QY 1300 GCGACCGTATGAGAT..... 1313
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
825 GCGACCGTATGAGATAGCAAGTCCATGCTGCGACACCAAGGTGATGCTGGATGAGCGAC 884
QY 1314 ..... 1325
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
885 GCGAAGTCTCTTCAAGACAGATTGATTATCTCTATCATCACTGCGAGGAGATGAAC 944
QY 1326 CAGAGTATACGATGCTGAGCATGAGAGACTGATCTTCAGGTGCTGCGCTGAGCGCCAGC 1385
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
945 CAGAGTATACGATGCTGAGCATGAGAGACTGATCTTCAGGTGCTGCGCTGAGCGCCAGC 1004
QY 1386 GCGCTCTTCAGGAGTCTATGCGAGTTCGCGAGCGCTCTTTTCATCATCACTGAGCGCTC 1445
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1385 GCGCTCTTCAGGAGTCTATGCGAGTTCGCGAGCGCTCTTTTCATCATCACTGAGCGCTC 1064
QY 1446 GATGAGTCTAGAGTCTTTTCCAGGATCTGAGGACCGCTGCTGCTGCTGCTGAGGAG 1505
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1065 GATGAGTCTAGAGTCTTTTCCAGGATCTGAGGACCGCTGCTGCTGCTGCTGAGGAG 1124
QY 1506 AAACGGCTAGGAGTGTGCTTTTACAGCTTAATTCGGAAGAAGCTGCTCTCTGAGCTA 1565
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1125 AAACGGCTAGGAGTGTGCTTTTACAGCTTAATTCGGAAGAAGCTGCTCTCTGAGCTA 1184
QY 1566 TCTTGTCTATATACAGCGGTCAGCGGTTTGGAGAACTCCACCGTGTCTGCTGAGGAC 1625
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1185 TCTTGTCTATATACAGCGGTCAGCGGTTTGGAGAACTCCACCGTGTCTGCTGAGGAC 1244
QY 1626 CCGAGGATGTGAGATCTGTGGCTTCTCTGAGGACAAAGGAAGTAATCTTCTACAG 1685
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1245 CCGAGGATGTGAGATCTGTGGCTTCTCTGAGGACAAAGGAAGTAATCTTCTACAG 1304
QY 1686 TATTTCAGATATCAAGAGAGGGGCTTAAGTTCATTAATTACGTGAGGACAAAGAGCT 1745

```

```

Db 1325 TATTTCACAAATGCGAGACAGCGGGCCCAAGTCTTCAATTACGTGAGGAGCAACAGAGCCT 1364
Qy 1745 CTCTTCACATGTGTCTGTCCTCCCTGTGTGCTGGTGGTGTGTACTGCTGCTCCAGAG 1805
Db 1365 CTCTTCACATGTGTCTGTCCTCCCTGTGTGCTGGTGGTGTGTACTGCTGCTCCAGAG 1424
Qy 1806 CAGCTGGAGGGGTGGGGGCTGTGAGACAGAGCTGTCCAGAGACAGCACTGCACTGTACATG 1865
Db 1425 CAGCTGGAGGGGTGGGGGCTGTGAGACAGAGCTGTCCAGAGACCACTGCACTGTACATG 1484
Qy 1866 CTCTACTCTGTAGTGTGTATGAACTGAAAGCGGCGGAGCGGAGGAGGAGGAGGAGGAGG 1925
Db 1485 CTCTACTCTGTAGTGTGTATGAACTGAAAGCGGCGGAGCGGAGGAGGAGGAGGAGGAGG 1544
Qy 1926 AACCAGAGAGGGGTGTGTCTGCTGTGGGCGAGATGAGTCTTGAATCGAAGAAATCTATT 1985
Db 1545 AACCAGAGAGGGGTGTGTCTGCTGTGGGCGAGATGAGTCTTGAATCGAAGAAATCTATT 1604
Qy 1986 GAGGAGCAGAGACTCCGGAAGACAGCGGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAG 2045
Db 1605 GAGGAGCAGAGACTCCGGAAGACAGCGGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAG 1664
Qy 2046 ATGACATCTTCCAGAGAGACATGAACTGTGAG 2078
Db 1665 ATGACATCTTCCAGAGAGACATGAACTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1724
Qy 2078 2078
Db 1725 TTCCAGAGAACTTTTTCGACATATCTACTATATATATATATATATATATATATATATAT 1784
Qy 2079 2079
Db 1785 CCAGAGCAGAGAGCTGACAGAGCTGTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1844
Qy 2091 GCACTCAGAGAGAGCTTCTCTGTATATATATATATATATATATATATATATATATATATAT 2150
Db 1845 GCACTCAGAGAGAGCTTCTCTGTATATATATATATATATATATATATATATATATATATAT 1904
Qy 2151 AAGAGCTGTCTGAAAGCTGTGCGGACATATAGATGAGAGAGAGAGAGAGAGAGAGAGAG 2210
Db 1905 AAGAGCTGTCTGAAAGCTGTGCGGACATATAGATGAGAGAGAGAGAGAGAGAGAGAGAG 1964
Qy 2211 AACAAGCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2270
Db 1965 AACAAGCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2024
Qy 2271 TTGTACAGATCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2330
Db 2025 TTGTACAGATCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2084
Qy 2331 GGTGTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2390
Db 2085 GTGTGTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2144
Qy 2391 TCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2450
Db 2145 TCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2204
Qy 2451 CGCGCAGAGTGTCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2510
Db 2205 CGCGCAGAGTGTCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2264
Qy 2511 CTGCTGGAAGCCTACAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2570
Db 2265 CTGCTGGAAGCCTACAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2324
Qy 2571 GAGCTGTCTGTACCGAATGCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2630
Db 2325 GAGCTGTCTGTACCGAATGCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2384
Qy 2631 CTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2690

```

```

Db 2385 CTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2418
Qy 2691 TATTTTGAATAAATATATTGGCCAGGTATATGATGCTCAGCCCTGTAATTCAGACATTT 2750
Db 2419 2419
Qy 2751 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2810
Db 2419 2419
Qy 2811 GGTGAATGCTATCTTACTATATATATATATATATATATATATATATATATATATATAT 2870
Db 2419 2419
Qy 2871 TAAGCCCACTACTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2930
Db 2419 2419
Qy 2931 TGTGCTTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2990
Db 2419 GGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2475
Qy 2991 CAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3050
Db 2475 CAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2515
Qy 3051 TGATCTGCTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3110
Db 2515 TGATCTGCTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2595
Qy 3111 GGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3170
Db 2595 GGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2655
Qy 3171 CAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3230
Db 2655 CAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2715
Qy 3231 TATGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3290
Db 2715 TATGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2775
Qy 3291 GGTGCTTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3350
Db 2775 GGTGCTTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2835
Qy 3351 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3410
Db 2835 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2895
Qy 3411 GGTGCTTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3470
Db 2895 GGTGCTTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2939

```

RESULT 4

AK095460

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

AK095460 3456 bp -RNA linear PRI 15-JUL 2002
 Homo sapiens cDNA FLJ38141 fis, clone D80ST2002673, weakly similar
 to Homo sapiens caspase recruitment domain protein 7 mRNA.
 AK095460.1 GI:21754727
 Cligo capping; fis (full insert sequence).
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 Ninomiya,K., Waga, S., M., Kanda,K., Kondo,H., Yokoi,T.,
 Kodaira,H., Furuya,T., Takahashi,M., Kikkawa,E., Omura,Y., Abe,K.,
 Kamihara,K., Katsuta,N., Sato,K., Tanikawa,M., Yamazaki,M.,
 Sugiyama,T., Irie,P., Otsuki,T., Sato,H., Nakamatsu,A., Ishii,S.,
 Yamaoto,O., Isomoto,Y., Kawai,H., Saito,K., Nishikawa,T.,

[illegible]

Db	2864	GCUTGAGCATCCAGCCTGCAAGACCTGAGACCTGGGCGTGGAGAGCTGTGGCTCACAG	292
Qy	3471	SCTAGG 3476	
Db	2924	CCAAAGG 2929	
RESULT 5			
BC028069			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REMARK			
COMMENT			

1548 AAGAGTCTCTGCTGGAAGTCTGCGGCACATCAAGATGACCTGTGCGAGTGCATCCAA 1907
 2221 AGCAAGCTGAGCGAGCGGTCTCAGCTTCAGCTTCAGAGAGTCTCTGAGTCTTCAGCTGC 2270
 1908 AGCAAGCTGAGCGAGCGGTCTCAGCTTCAGCTTCAGAGAGTCTCTGAGTCTTCAGCTGC 1967
 2271 TTGTACAGATCCAGGAGGAGGAGTTTATCAAGCAAGCTTCAGGAGCTTCAGGAGTGC 2330
 1968 TTGTACAGATCCAGGAGGAGGAGTTTATCAAGCAAGCTTCAGGAGCTTCAGGAGTGC 2027
 2331 TTGTACAGATCCAGGAGGAGGAGTTTATCAAGCAAGCTTCAGGAGCTTCAGGAGTGC 2390
 2028 TTGTACAGATCCAGGAGGAGGAGTTTATCAAGCAAGCTTCAGGAGCTTCAGGAGTGC 2067
 2431 TGTAGGAGCGCGCGAGGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2450
 2099 TGTAGGAGCGCGCGAGGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2147
 2451 CGCGGAGGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2510
 2148 CGCGGAGGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2207
 2511 CTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2570
 2267 CTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2267
 2571 GAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2630
 2327 GAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2327
 2631 CTGAGACCGCCAACTGCAAACTGCAAACTGCAAACTGCAAACTGCAAACTGCAAACTGCA 2690
 2428 CTGAGACCGCCAACTGCAAACTGCAAACTGCAAACTGCAAACTGCAAACTGCAAACTGCA 2361
 2691 TATTTTGAAATAATATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2750
 2362 TATTTTGAAATAATATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2361
 2751 GCGAGCGCGAGTGGGAGGAGTGGGAGGAGTGGGAGGAGTGGGAGGAGTGGGAGGAGTGGG 2810
 2362 GCGAGCGCGAGTGGGAGGAGTGGGAGGAGTGGGAGGAGTGGGAGGAGTGGGAGGAGTGGG 2361
 2811 GGTAAAGCTGCTCTTAAATAATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2870
 2362 GGTAAAGCTGCTCTTAAATAATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2361
 2871 TAAGCGGAGTCTGAGGAGGAGTCTGAGGAGGAGTCTGAGGAGGAGTCTGAGGAGGAGT 2930
 2362 TAAGCGGAGTCTGAGGAGGAGTCTGAGGAGGAGTCTGAGGAGGAGTCTGAGGAGGAGT 2361
 2931 TTGTGCTGAGGAGTGGGAGGAGTGGGAGGAGTGGGAGGAGTGGGAGGAGTGGGAGGAGT 2990
 2362 TTGTGCTGAGGAGTGGGAGGAGTGGGAGGAGTGGGAGGAGTGGGAGGAGTGGGAGGAGT 2361
 2991 TGGCAATAAGAAATTTGACAGAGTGGTCTGATGAGGAGGAGTGGTCTGATGAGGAGGAG 3050
 2419 TGGCAATAAGAAATTTGACAGAGTGGTCTGATGAGGAGGAGTGGTCTGATGAGGAGGAG 2478
 3051 TGGATGCTGCTTGGAGGAGTGGGAGGAGTGGGAGGAGTGGGAGGAGTGGGAGGAGTGGG 3110
 2479 TGGATGCTGCTTGGAGGAGTGGGAGGAGTGGGAGGAGTGGGAGGAGTGGGAGGAGTGGG 2538
 3111 GGAAGTGTGAGTGTGAGGAGTGGGAGGAGTGGGAGGAGTGGGAGGAGTGGGAGGAGTGGG 3170
 2539 GGAAGTGTGAGTGTGAGGAGTGGGAGGAGTGGGAGGAGTGGGAGGAGTGGGAGGAGTGGG 2598
 3171 CACATCTGAGTGTGAGTGTGAGGAGTGGGAGGAGTGGGAGGAGTGGGAGGAGTGGGAGG 3230
 2599 CACATCTGAGTGTGAGTGTGAGGAGTGGGAGGAGTGGGAGGAGTGGGAGGAGTGGGAGG 2658
 3231 TATGCGGAGGAGTGGGAGGAGTGGGAGGAGTGGGAGGAGTGGGAGGAGTGGGAGGAGTGGG 3290

2659 TATGCGGAGGAGTGGGAGGAGTGGGAGGAGTGGGAGGAGTGGGAGGAGTGGGAGGAGTGGG 2718
 3291 GCGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3350
 2719 GCGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2778
 3351 GAGAGCTGAGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 3410
 2779 GAGAGCTGAGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 2838
 3411 GCGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3454
 2819 GCGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2882

RESULT 6
 AV116205
 LOCUS Homo sapiens monarch-1 splice form 11 mRNA, complete cds.
 DEFINITION alternatively spliced.
 ACCESSION AV116205
 VERSION AV116205.1 GI:21711822
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 1563)
 AUTHORS Williams K.L., Takano D., Linhoff M.W., Reed W. and Ling J.P.Y.
 TITLE Monarch-1: A Pyrimidine-Binding Domain/Leucine-Rich Repeat Protein That Controls Classical and Nonclassical MHC Class I Genes
 JOURNAL C. Immunol. 173 (11), 5354-5359 (2003)
 PUBMED 12759429
 REFERENCE 2 (bases 1 to 1563)
 AUTHORS Williams K.L., Linhoff M.W. and Ling J.P.Y.
 TITLE Direct Submission
 JOURNAL Submitted (09-MAY-2002) Lineberger Cancer Center, UNC, Mason Farm Road, Chapel Hill, NC 27599, USA
 FEATURES
 source Location/Qualifiers
 1..1563
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /chromosome="19"
 /map="19q13.4"
 /cell_line="H937"
 221..3241
 /note="contains alternatively spliced exon 9"
 CDS
 3241..3241
 /codon_start=1
 /product="monarch-1 splice form 11"
 /protein_id="AA075143.1"
 /gb_xref="GI:21711822"
 /translation="MPLTAGDGGCPLSTYDEEAVECKKFKLYSTATELGEKIP
 WSKKAGRIEMAGLILTFEPPEARLALSTFERINRKLMDERNGREDLVDRTPRG
 PSSLGNGSTGLLESLVTPKQDQYRDVVRKFRIMEDRNARLGEVNLSHRYRL
 LKPKHNPVQOOLDTGRHARTVGHQASPKIETLEPDEEPERPEPTVYVGA
 AGICSLAKVRLDWDADKLPQGRDYLYINCRMNQATSCQDLIFSCWPPS
 APLGSLRVERLLFTLDGDEKPSFDQGWPCWCEERKPELILSLNRKLLP
 ELSLTTRFTALKRHLRHPHVEILGFSEARKEYKRYFHANQGVENYVR
 DNEPLTFMCPVPLVWVVTCTCQQLGGGLLRQTSRTTITAVYVLYLLSYQKPGAP
 KIQPEPNRGLCSLAADGLANOKILPEEDLRKHLGDGEDVSAPLNNIPKQINGFR
 YYSFHLSEFFFAAYVYLLDEGGAGDPDDVTRLTEVAFSESRFLATLSEPLFGL
 LNETRSHLRSKLCWVSRPHIKMLLQKSKAGSDGSLQSGLEFFSCLVEIQEE
 FIQAGSHQVQVYSNTRSMHNVSSFLKRCRSQVILHYATGATSGAGENRCSA
 GAHTLVGLDPERVTLDAYSHELAALCTNPNLIELSLYRNALSGRGLKQCSYR
 PCKLQRLKRCRISISSACEDLSAALANKNLTRMDLSNGVGFPMVLLCEGRHP
 QCKLQVQLSKCOLESACQEVSLGTNPVHVELDTGNALDELGLRLCCGLRHP
 CRLSTLKLICRLTAACDELASTLSVNQSLRELDLSLNLGLDGLVLLCEGLRHP
 KLCRLGLRGLSGACQGLSVYLNQHNRLBELDLSFNDLGDWGLLAELSGJHPACR
 LCKLWLLFGNELNKTHTSRJALRVTKPYLDIC"
 869..992
 /note="Region: P-loop"
 2059..3238
 misc_feature
 2059..3238
 misc_feature

[illegible]

QY	3129	CGGGGCTTCTCAGGASATGGGCTTCTGTGCTTGGACCAACCCACATCTCGTTGAGTTGG	1189
DB	2360	CGGGGCTTCTCAGGASATGGGCTTCTGTGCTTGGACCAACCCACATCTCGTTGAGTTGG	2419
QY	3189	ACCTGACAGCAATCACTGAGGATTTGGGCTGAGGTACTATGCGCAGGAGCTAGGC	3248
DB	2420	ACCTGACAGCAATCACTGAGGATTTGGGCTGAGGTACTATGCGCAGGAGCTAGGC	2479
QY	3249	ACGACAGCTTCTGAGCACTGCAATTTTGGCTTGAAGATCTGGCGCTCATCTGCTGCT	3309
DB	2480	ACGACAGCTTCTGAGCACTGCAATTTTGGCTTGAAGATCTGGCGCTCATCTGCTGCT	2539
QY	3309	GTGACAGCTTCTGAGCACTGCAATTTTGGCTTGAAGATCTGGAGCTGAGAGCTTGAGCT	3369
DB	2540	GTGACAGCTTCTGAGCACTGCAATTTTGGCTTGAAGATCTGGAGCTGAGAGCTTGAGCT	2599
QY	3369	TGATGAGCTTGGGCACTGCGGCTTCTGTGCTGAGGCTTCAGGCATTCACCT	3428
DB	2600	TGATGAGCTTGGGCACTGCGGCTTCTGTGCTGAGGCTTCAGGCATTCACCT	2659
QY	3429	GCAAGCTGAGAGCTTGGCTGAGCT	3454
DB	2660	GCAAGCTGAGAGCTTGGCTGAGCT	2685
RESULT 9	AX459869	LOCUS	AX459869
DEFINITION	Sequence 12 from Patent WO240665.	DNA	3400 bp
ACCESSION	AX459869	Linear	PAT 98-JUL-2002
VERSION	AX459869.1		
KEYWORDS	Sequence 12 from Patent WO240665.		
SOURCE	Human sapiens (human)		
ORGANISM	Human sapiens		
REFERENCE	Exkabyov, Mrazov, Chodara, Vertebrata, Euteleostomi;		
AUTHORS	Mammalia, Euteleostomi, Vertebrata, Euteleostomi;		
TITLE	1		
JOURNAL	Proteins and dna sequences underlying these proteins used for		
FEATURES	treating inflammation		
SOURCE	Patent: WO 0240668 A 12 23 MAY 2002		
	Apotech Research and Development Ltd. (CH)		
	Location/Qualifiers		
	1..3320		
	Organism="Homo sapiens"		
	mol_type="genomic DNA"		
	date_xref="taxon:9606"		
	ref="MIM:104310" /		
	126..343		
	652..652		
BASE COUNT	126..343		
ORIGIN	126..343		
Query Match	32.4%	Score 1416;	DB 6; Length 3300;
Best local Similarity	41.8%	Prod No. 0;	
Matches 2160;	Conservative	0; Mismatches	5; Indels 581; Gaps 4;
QY	1021	GATCCGCAAGAAATATGAGGAAATATGTCGACAGAAATTCGCGCTCATGAGACCGC	1080
DB	289	GATCCGCAAGAAATATGAGGAAATATGTCGACAGAAATTCGCGCTCATGAGACCGC	349
QY	1081	AATGGCGCTATGAGAAATGTCTCAACTTCAGCAACCGGTACACCGGCTCTGTGTGGT	1140
DB	349	AATGGCGCTATGAGAAATGTCTCAACTTCAGCAACCGGTACACCGGCTCTGTGTGGT	408
QY	1141	AAGGAGCACTCAAACTTATGAGCTTCAGCTTCAGCAGCAGCTTCTGGACACAGGCGGGAAC	1200
DB	409	AAGGAGCACTCAAACTTATGAGCTTCAGCTTCAGCAGCAGCTTCTGGACACAGGCGGGAAC	468
QY	1201	GTCAGAGCTTGGGACACACAGCTTACCGCTTCAAGATAGAGACCTCTTTGAGCCAGAC	1260
DB	469	GTCAGAGCTTGGGACACACAGCTTACCGCTTCAAGATAGAGACCTCTTTGAGCCAGAC	528
QY	1261	GTCAGAGCTTGGGACACACAGCTTACCGCTTCAAGATAGAGACCTCTTTGAGCCAGAC	1320

FEATURES		Location/Qualifiers	
SOURCE	1..3221	/organism="Homo sapiens"	
		/mol_type="mRNA"	
		/db_xref="taxon:3606"	
		/chrchromosome="19"	
		/map="19q13.4"	
CDS	221..3299	/note="contains alternatively spliced exons 7, 8 and 9"	
		/codon_start=1	
		/product="neonarchin splice for 1V"	
		/protein_id="AA075451"	
		/db_xref="GI:21911827"	
		/translation:"MURTAGDGGGCHLSTYLEAEVAKKFLDYLTATELGEKIP WGVKAGPJEKACLIITHPGPEAMRLASTFERINRKLWREGREGLVROTPEGG PSSGNGSLCLRLSVLTPKPKQETVRYVPAKPMEDORNAELCEVKNLSHYRRL LLVKEHNPVQVQVLLDTPGHARTVCHOASPIKLETFEPDEEPERPRVVMWGA AGTCKSLAHKMLDRAGKSLFGKRFKLYINCREYNQSRATEGSDQDILFESWPEFS APICELLIRVERLJFTGGDELKPSHPDQGFMCWMBEERPELELNSLJHKKLLP ELSLIITRETALEKLRLEHERHVEELGSEAKRKFKYKYPKRNABAGQVFNVR DNBPFTMCPEVLVWVCTCLOQLEGGSLRQTSRTTAVVYMLYLLSLMOPKGPAP RLQPPNORNGICSLAAGLANQILFBEODLPKGGIDGEDVSAPLANNIFPKQINER VYSEFHLSPQEEFAAMYLLDEEGGAGPQVLPRLTEVAFSEPSFLALTSRLKGL LASEITSHLEKSLCWKVSRLKMDLQWKAQNSGSLTQCSLEFFSLCLVIOBEE FQCALSHFVIVVSN:ANKMENVSSFCUKRCSACVJHEUATISAGDEURARSA GATLLVQKPERVILCAYSERLAALCTNPVLELSVFNALGSGRKLKLCQGLRH PKLQNGRLKRLRISSEAGELAAALIAKNIKPYLSINQVGFPMKILCRGRHP QCRILQMLPKCKQLESGATCEMASVLSTNEHLVEELDTGNALDQGLRLGCGTLRHPV CRLLTLWLPFGMDLNKNTHSRLAALPVTRFVYDLC"	
	misc_feature	669..892	/note="Region: P loop"
	misc_feature	2369..2896	/note="Region: leucine rich repeat"
	misc_feature	221..289	/note="Region: P/kin domain"
	BASE COUNT	749 a 891 c 336 g 645 t	
	ORIGIN		
Query Match			
Best local similarity 85.4%; Pident. No. 0.			
Values 2667; Conservative 0; Mismatches 2; Indels 8 501; Gaps 4;			
QY	940	GATACCCGACCTGTTGCGGCTGCTTACTTGGAGACATTTAAATGCGCTTGTGAAGTC	999
DB	509	GATACCCGACCTGTTGCGGCTGCTTACTTGGAGACATTTAAATGCGCTTGTGAAGTC	568
QY	1000	TCTTCTGCTACTCCAAAGAAAGATGACAAAGCTTAAAGTAAAGTATGCTGATAGMAA	1059
DB	569	TCTTCTGCTACTCCAAAGAAAGATGACAAAGCTTAAAGTAAAGTATGCTGATAGMAA	629
QY	1260	TTGGGCTGATGGAAGACCGCAATATGCTGATATGATATGATATGATATGATATGAT	1319
DB	629	TTGGGCTGATGGAAGACCGCAATATGCTGATATGATATGATATGATATGATATGAT	688
QY	2120	TACAGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2179
DB	689	TACAGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	748
QY	2180	CTGACACAGCGCTGGACACGCTGACACGCTGACACGCTGACACGCTGACACGCTGAC	2239
DB	749	CTGACACAGCGCTGGACACGCTGACACGCTGACACGCTGACACGCTGACACGCTGAC	808
QY	1240	GAGACCTCTTTTGGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC	1299
DB	809	GAGACCTCTTTTGGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC	868
QY	1300	GGCGCGCAGGGAT.....	1313
DB	869	GGCGCGCAGGGATAGGCAAT	928
QY	1314	1325
DB	929	GGGAGAGCTCTTCCAGGACGACATTTATATATATATATATATATATATATATATATAT	988

Db 223 ACCAGGACCCACCTGGAGAGAGCTCTGTCTGGAAGGTCCTGGCGACATCAGATGGAC 192
 QY 2191 CTGTGTGAGTGGATCCAAAGCAAGCTCAGAGTACAGGCTCCACCGCTCAGGAGGCTCC 2252
 Db 284 CTGTGTGAGTGGATCCAAAGCAAGCTCAGAGTACAGGCTCCACCGCTCAGGAGGCTCC 342
 QY 2253 TTGAGTGTCTTCAGTCTCTGTATATATATATATATATATATATATATATATATATAT 2312
 Db 342 TTGAGTGTCTTCAGTCTCTGTATATATATATATATATATATATATATATATATATAT 402
 QY 2333 AGCTACCTTCAGGTGATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2372
 Db 453 AGCCACCTTCAGGTGATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 462
 QY 2373 TCGTCTGTCTGTAGAGGCTGTAGAGGCTGTAGAGGCTGTAGAGGCTGTAGAGGCTGTAG 2432
 Db 463 TCGTCTGTCTGTAGAGGCTGTAGAGGCTGTAGAGGCTGTAGAGGCTGTAGAGGCTGTAG 522
 QY 2433 AGCCGGAGGGGAGAGCCGGGCGGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2492
 Db 523 AGCCGGAGGGGAGAGCCGGGCGGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 582
 QY 2453 CTACAGAGAGGAGCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2552
 Db 584 CTACAGAGAGGAGCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 642
 QY 2553 ACCAATCCAAACCTGTATAGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2612
 Db 643 ACCAATCCAAACCTGTATAGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 702
 QY 2613 AAGTGTCTGTCAAGGACTCAGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2672
 Db 703 AAGTGTCTGTCAAGGACTCAGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 754
 QY 2673 TTATCATATACATGATATTTTAAATTAATATATCTGGTAACTATATGTGTACGCG 2732
 Db 755 754
 QY 2733 CTGTAAATCCAGGACTTTGAGGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2792
 Db 755 754
 QY 2793 GACGAGCTGCCCACATAGCTCAAAAGGATCTCTTAATAAATACCAAAATGCGCAGG 2852
 Db 755 754
 QY 2853 CATGTTGACACACCTGTCTAAGCTTATATATATATATATATATATATATATATATAT 2912
 Db 755 754
 QY 2913 AACTCAGAGGACAGAGCTGT 2972
 Db 755 754
 QY 2973 ACCCTCTCGAGCTCTCATAGCCCAATATATATATATATATATATATATATATATATAT 3032
 Db 794 ACCCTCTCGAGCTCTCATAGCCCAATATATATATATATATATATATATATATATATAT 853
 QY 3033 GCGTTGATCTCCAGGACTGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3092
 Db 854 GCGTTGATCTCCAGGACTGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 913
 QY 3093 TGCACATGATTCAGTTAGGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3152
 Db 914 TGCACATGATTCAGTTAGGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 973
 QY 3153 CTGTCTCTGGACCAACCCACATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3212
 Db 974 CTGTCTCTGGACCAACCCACATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1033
 QY 3213 ATTTGGGCTGTAGGTTACTATGTGCAAGGACTGAGGAGTGTGTGTGTGTGTGTGTGTGTGT 3272

Db 1034 ATTGGGCTGAGGTTATATATGACAGGAGTGAAGGACCCAGCTGTGAGACTACGGAATT 1093
 QY 3273 TGTGCTGAGATATGCGGCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3332
 Db 1094 TGTGCTGAGATATGCGGCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1153
 QY 3333 GTGTGAGACAAGCTTGAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 3392
 Db 1154 GTGTGAGACAAGCTTGAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1213
 QY 3393 TGTGCTGAGATATGCGGCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3452
 Db 1214 TGTGCTGAGATATGCGGCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1273
 QY 3453 TC 3454
 Db 1274 GC 1275

RESULT 14
 AF231021
 LOCUS Homo sapiens leucine rich-repeat protein RNO2 mRNA, complete cds.
 DEFINITION AF231021
 ACCESSION AF231021
 VERSION AF231021.1 GI:13318296
 KEYWORDS
 SOURCE Homo sapiens (human).
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 2494)
 AUTHORS Sharif, P. C., Khandel, N., Wang, L. Y., Vreeke, T. M. and Parker, C. J.
 TITLE Identification and characterization of a novel gene that is
 upregulated in leukemia cells by nitric oxide
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 2494)
 AUTHORS Shami, P. J., Khandel, N., Wang, L. Y., Vreeke, T. M. and Parker, C. J.
 TITLE Direct Sub-Region
 JOURNAL Submitted (02-FEB-2000) Medicine, University of Utah and SLC VA
 Medical Centers, Box 151M, 500 Foothill Boulevard, Salt Lake City,
 UT 84143, USA

FEATURES
 Location/Qualifiers
 1..2494
 /organism="Homo sapiens"
 /mol_type="RNA"
 /db_xref="taxon:9606"
 1309..2192
 /note="Similar to the Homo sapiens ribonuclease/angiogenin
 inhibitor, expressed in hematopoietic cells; upregulated
 by nitric oxide in leukemia cells"
 /codon_start=1
 /product="leucine-rich-repeat protein RNO2"
 /protein_id="AAK4942.1"
 /db_xref="GI:13318297"
 /translation="MSCAMWHTSVSPATQAKAGLLPQRRQLWLRKRRISSACRDL
 LSAALANKLITRMDLGSNGVPGMVLCEGLRHPQCRLOMQLKRCOLSCACQEM
 ASVGLNPHVLELDLTGNALDELGLRLLCGGLRHPVCRLETLKICRUTAAACDLA
 STLVNGLRELDLSNELSDGLLCEGLRHPVCRLETLKICRUTAAACDLA
 TSVNGLRELDLSNELSDGLLCEGLRHPVCRLETLKICRUTAAACDLA
 RVTPKYDGGC"

BASE COUNT 634 a 636 c 699 g 525 t
 ORIGIN
 Query Match 15.9%; Score 183.6; DB 9; Length 2494;
 Best Local Similarity 90.0%; Pred. No. 4e-170;
 Matches 884; Conservative 0; Mis-matches 19; Indels 79; Gaps 1;

QY 2495 ACCAGAGGAGCGCTTCTGTGGAGCGCTACAGTGAACATCTGCAGCGCCCTGTGCAC 2554
 Db 1043 ACCAGAGGAGCGCTTCTGTGGAGCGCTACAGTGAACATCTGCAGCGCCCTGTGCAC 1093
 QY 2555 GATCCAAAGCTGATAGAGCTGTCTGTATGGAATGCCCTGCGAGCGCGGAGGTGAA 2614

1100 CAATCCAAATCTGATAGAGCTGTCTCTCTACAGAAATGCTCTGGACAGCGGGGGGTAA 1150
11615 GTTGCTCTGTCAAGGACTCAGACATCCCAACTGCAGAGCTTCAGAGCTCAGAGCTTAAATTT 2674
1160 GTTGCTCTGTCAAGGACTCAGACATCCCAACTGCAGAGCTTCAGAGCTTAAATTT 1208
12675 ATCATATATAATGATATTTTGAAATATAATATATTTGGTAATGATATGCTCAGGCT 2734
1209 ----AGACATTTTGGAGTTTATATATATATATATATATATATATATATATATATAT 1209
2145 GTAATTCAGACACTTGGTAGGTTTAAATTTAGAGATATATATATATATATATATATAT 2794
1209 ----AGACATTTTGGAGTTTATATATATATATATATATATATATATATATATATAT 1260
2295 ATAGCTGGCCACATCTTGAAAGCTGATCTTATATATAAATATTAATATAGCTAGGTA 1854
1261 ATAGCTGGCCACATCTTGAAAGCTGATCTTATATATAAATATTAATATAGCTAGGTA 1320
1855 TGGTGACACAGCTGTATAGCTGATCTTATATATAAATATTAATATAGCTAGGTA 2914
1321 TGGTGACACAGCTGTATAGCTGATCTTATATATAAATATTAATATAGCTAGGTA 1380
1915 CCGAGAGGACAGAGTGTGTGGCTGAAAGATTTATATATATATATATATATATATATAT 2974
1361 CCGAGAGGACAGAGTGTGTGGCTGAAAGATTTATATATATATATATATATATATATAT 1440
2375 CTCTCTGGAGCTCTATAGCCATTAAGATTTTAAAGATATGATCTGATGTCAGAGGCT 4034
1441 CTCTCTGGAGCTCTATAGCCATTAAGATTTTAAAGATATGATCTGATGTCAGAGGCT 1500
1636 GTTGGATTCACAGCATGATGCTGCTTTTGAGAGGCTTGTGATGCTTCTATAGAGGCTG 3094
1501 GTTGGATTCACAGCATGATGCTGCTTTTGAGAGGCTTGTGATGCTTCTATAGAGGCTG 1560
1695 CAGATGATTCAGTTCAGAGAGTGTGATGATGCTGCTGCTTGTGATGCTGCTGCTGCT 3154
1561 CAGATGATTCAGTTCAGAGAGTGTGATGATGCTGCTGCTTGTGATGCTGCTGCTGCT 1620
1755 GTCTTGGACCAACCATCATCTGTTGATGATGCTGCTTGTGATGCTGCTGCTGCTGCT 3214
1621 GTCTTGGACCAACCATCATCTGTTGATGATGCTGCTTGTGATGCTGCTGCTGCTGCT 1680
1815 TTGGGCTGAGTTTACTATGCTGAGAGTGTGATGATGCTGCTTGTGATGCTGCTGCTGCT 3274
1681 TTGGGCTGAGTTTACTATGCTGAGAGTGTGATGATGCTGCTTGTGATGCTGCTGCTGCT 1740
1875 TGGCTGAGATCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3334
1741 TGGCTGAGATCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800
1915 GTGAACAGAGGCTTGAGAGAGTGTGATGATGCTGCTTGTGATGCTGCTGCTGCTGCT 3394
1801 GTGAACAGAGGCTTGAGAGAGTGTGATGATGCTGCTTGTGATGCTGCTGCTGCTGCT 1860
1995 CTGCTGCTGTGTGAGGCTTCAGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3454
1861 CTGCTGCTGTGTGAGGCTTCAGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1920
3455 CCGTTTTCCTCACCAGCTAGG 3476
1921 AGCTGTGGCTTCACAGCCAAG 1942

RESULT 15
AC156193
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AC106193
Rattus norvegicus clone CH210 259820, WORKING DRAFT SEQUENCE.
AC106193
AC106193.4 G130579353
HTG: HTGS_PHASES; HTGS_DRAFT; HTGS_FULLTOP.
Rattus norvegicus (Norway rat)
Eukaryota; Metazoa; Chordata; Glires; Euteleostomi;

223144 bp LNA
HTG 13-MAY-2003

REFERENCE	TITLE
AUTHORS	JOURNAL
	REFERENCE
	AUTHORS
	TITLE
	JOURNAL
	REFERENCE
	AUTHORS
	TITLE
	JOURNAL
	COMMENT

Mammalia; Eutheria, Rodentia; Sclerognathi; Muridae; Murinae;
Rattus.

1 Cases 1 to 223344:
Muzny,D,Wylie,M, Metzger,K, Lee,S, Abrazon,S, Adams,C, Alder,J,
Alder,C, Allen,H, Alstonsen,S, Akin,A, Arguano,D,
Anyatebechi,V, Ayadi,X, Ayodeji,M, Baca,E, Baden,H,
Baldwin,D, Bandaranake,D, Barber,M, Barnstead,M, Behrmann,F,
Bedwin,K, Blair,B, Blankenship,K, Byth,P, Brown,M,
Pryant,N, Buiay,C, Huchra,J, Burrell,K, Calderon,E,
Cardenas-V, Carter,K, Cavazos,I, Cessari,H, Center,A,
Chacón,G, Chavez,D, Chen,S, Chen,R, Chen,Y, Chen,Z, Chu,J,
Cleveland,K, Cockrell,B, Cox,C, Coyle,M, Cree,A, O'Shuara,
David,L,M, Davis,C, Davy Carrillo,J, De Anda,C, Dederich,D,
Delgado,O, Denison,S, Deramo,C, Ding,Y, Dinh,R, Divya,K,
Drapar,H, Duran-Fochia,S, Dunn,A, Durbin,K, Duval,B, Evans,K,
Egan,A, Escoto,M, Eugene,C, Evans,C,A, Falls,J, Fan,G,
Farrand,S, Fawcett,M, Feilley,M, Flagin,N, Forbes,D, Foster,M, Foster,P,
Frederick,C,V, Gabler,A, Ganta,R, Garcia,A, Garner,T, Garza,M,
Feaser,C, Geier,K, Gill,R, Grady,M, Guerra,T, Guzman,W,
Gubregovigis,E, Geor,K, Gili,R, Glick,C, Hamilton,C, Hamilton,K,
Gurathane,P, Haaland,W, Harris,C, Henderson,N, Hernandez,S,
Harvey,J, Havak,P, Hawes,A, Henderson,N, Hernandez,S,
Hernandez,R, Hines,S, Hladik,S,B, Hodgson,A, Hoques,M,
Hollins,P, Howell,S, S, Huily,S, Hume,J, Idler,D,D, Jackson,A,
Jackson,L, Jacob,B, Johnson,B, Johnson,D, Johnson,R, Colivet,A,
Kachapathy,S, Kelly,S, Kelly,S, Khan,Z, King,S, Kovac,C,
Kowals,C, Kraft,C,L, Lebow,H, Levani,J, Lewis,L, Liu,Z, Liu,C,
Liu,J, Liu,W, Liu,Y, Lordon,P, Longacre,S, Lopez,J,
Lukacs-Szawal,V, Lunegrad,H, Lozada,R,J, Lu,X, Ma,J,
Maheshwari,V, Mahindaratne,K, Mahmood,M, Malloy,K, Mangum,A,
Manjuri,S, Mapa,P, Martin,K, Martin,R, Martinez,E,
Mathewney,S, McLeod,X,F, McNeill,T,Z, Meenen,E,
McGowan,L, Mervin,A, Miner,G, Mirja,E, Montemayor,J, Moore,S,
Morghan,M, Morris,K, Morris,S, Mundasa,M, Murphy,M, Nair,B,
Markkervis,C, Neal,D, Newton,N, Nguyen,N, Norris,S,
Nwakileleh,O, Okunni,G, Olajunsgoon,A, Pal,S, Parks,K,
Pasternak,S, Paul,H, Perez,A, Perez,L, Plankoch,C,
Proffner,P, Pandexter,A, Popovic,D, Prius-E, Pub-L,
Puozzo,M, Ponzio,J, Rachlin,E, Reeves,K, Regier,M,A, Reid,R,
Reilly,B, Reilly,M, Ben,Y, Reiter,M, Richards,S, Riggs,F,
Rives,C, Rokey,T, Rojas,A, Rose,M, Rose,R, Ruiz,S,
Sanders,W, Savery,G, Scherer,S, Scott,G, Shatsman,S, Shen,H,
Shetty,J, Shivasteyn,A, Sisson,I, Sitter,C,D, Straz,D,
Sured,A, Sodergren,E, Song,X-Z, Sorrelier,R, Sosai,J,
Stearle,M, Strong,R, Sutton,A, Svatek,A, Tabot,P, Taylor,J,
Taylor,T, Thomas,N, Thomas,S, Tingey,A, Trejos,Z, Ugmani,K,
Valas,P, Vera,V, Villasantia,B, Waldron,L, Walker,B, Wang,C,
Wang,Q, Wang,S, Warren,J, Warren,K, Wei,X, White,F,
Williams,G, Wilson,R, Wiczysk,R, Wooden,H, Worley,K,
Wyck,D, Wright,S, Wu,J, Yakub,S, Yen,J, Yoon,S, Yoon,V,
Yu,P, Zhang,J, Zhou,Y, Zhou,X, Zhao,S, Dunn,D, von
Neudtshausen,A, Weiss,R, Smith,D.R., Holt,R.A., Smith,H.C.,
Weinstock,G, and Gibbs,F.A.
Direct Submission
Unpublished

2 (bases 1 to 223344)
Worley,K.C.

Direct Submission
Submitted 12-JAN-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 223344)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (13-MAY-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

On May 19, 2003 this sequence version replaced gi|223609300.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
<http://www.tisc.bcm.tmc.edu/projects/rat/>. Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a "contig-scaffold"). Within each contig-scaffold,
individuals' sequence contigs are ordered and oriented, and separated

by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

```

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GJ02
Center clone name: CH230-223344
----- Summary Statistics
Assembly program: Atlas 3.2
Consensus quality: 212092 bases at least Q40
Consensus quality: 214628 bases at least Q30
Consensus quality: 216512 bases at least Q20
Estimated insert size: 225123; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum of contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* See http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html.
* NOTE: This is a 'working draft' sequence. It currently
  consists of 1 contigs. Gaps between the contigs
  are represented as runs of N. The order of the pieces
  is believed to be correct as given, however the sizes
  of the gaps between them are based on estimates that have
  provided by the submitter.
  This sequence will be replaced
  by the finished sequence as soon as it is available and
  the accession number will be preserved.
* 223344: contig of 223344 bp in length.
  Location/Qualifiers
    id: 223344
    /organism="Rattus norvegicus"
    /mol_type="genomic DNA"
    /db_xref="taxon:10116"
    /clone="CH230 223344"
    990..1643
misc_feature
  /note="clone boundary"
  clone_end:17
  site:EcoRI
  end sequence:PWK(D:0756)
  77257..78553
misc_feature
  /note="wgs contig"
  219210..223047
misc_feature
  /note="clone boundary"
  clone_end:156
  site:EcoRI
  end sequence:PWK(D:0756)
  225799..223344
misc_feature
  /note="wgs end extension"
  clone_end:156
  58709 a 51139 c 42175 g 5484 t 5167 others

BASE COUNT
CPHIN

Query Match:
Best Local Similarity 69.08; Prod No. 560154;
Matches 1138; Conservative 0; Mismatches 377; Indels 208; Gaps 2;

QY 1020 AGATCCCGAGAACCTACAGGAGCTATGTCGAGAGAAATCTGTCATGAGAGAGCGG 1079
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
150339 AGATCCCGAGAACCTACAGGAGCTATGTCGAGAGAAATCTGTCATGAGAGAGCGG 150398
QY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1580 CAATCCCGGCTAGGGAATGTCACCTAGTCACCGGATATACCGGATACCGGCTGCTGT 1139
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
150399 CAATCCCGGCTAGGGAATGTCACCTAGTCACCGGATATACCGGATACCGGCTGCTGT 150458
QY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
150459 AAAGGAGACTCAATCTATCTAGGAGATATAGGAGATATAGGAGATATAGGAGATAT 150518
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

QY 1200 CCGGAGACCGTGGGACACACAGGCTAGCCCATTAAGATAGAGACCTCTTTGAGCCAGA 1259
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
150519 CCGGAGACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 150578
QY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1260 CAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1313
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
150519 TGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 150568
QY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1314 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
150639 GTCATATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCAT 150698
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
150639 GTCATATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCAT 150758
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1346 CATGAGAGCTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCAT 1405
QY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
150759 TGTACATGAGCTGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCAT 150818
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1406 CCGAGTTCGCGAGGCGCTCTTTTCATCATGCGAGCGTTCGATGAGTCAAGGCTTCCTT 1465
QY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
150819 CCGAGTTCGCGAGGCGCTCTTTTCATCATGCGAGCGTTCGATGAGTCAAGGCTTCCTT 150878
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1466 CCGAGTTCGCGAGGCGCTCTTTTCATCATGCGAGCGTTCGATGAGTCAAGGCTTCCTT 1525
QY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
150819 CCGAGTTCGCGAGGCGCTCTTTTCATCATGCGAGCGTTCGATGAGTCAAGGCTTCCTT 150938
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1526 TTTTAAAGAGTAAATGGGAGAGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1585
QY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
150939 CCGTGGAGAGCTGATTTGGAGGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 150998
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1586 CCGAGTTCGCGAGGCGCTCTTTTCATCATGCGAGCGTTCGATGAGTCAAGGCTTCCTT 1645
QY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
150939 CCGTGGAGAGCTGATTTGGAGGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 16058
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1646 CCGTGGAGAGCTGATTTGGAGGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1605
QY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
151059 CCGTGGAGAGCTGATTTGGAGGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 16118
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1706 CCGAGTTCGCGAGGCGCTCTTTTCATCATGCGAGCGTTCGATGAGTCAAGGCTTCCTT 1765
QY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
151119 CCGAGTTCGCGAGGCGCTCTTTTCATCATGCGAGCGTTCGATGAGTCAAGGCTTCCTT 151178
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1766 CCGAGTTCGCGAGGCGCTCTTTTCATCATGCGAGCGTTCGATGAGTCAAGGCTTCCTT 1825
QY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
151119 TCGATGAGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 151238
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1826 CCGAGTTCGCGAGGCGCTCTTTTCATCATGCGAGCGTTCGATGAGTCAAGGCTTCCTT 1885
QY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
151219 TTTTAAAGAGTAAATGGGAGAGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 151298
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1886 CCGAGTTCGCGAGGCGCTCTTTTCATCATGCGAGCGTTCGATGAGTCAAGGCTTCCTT 1945
QY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
151299 CCGAGTTCGCGAGGCGCTCTTTTCATCATGCGAGCGTTCGATGAGTCAAGGCTTCCTT 151358
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1946 CCGAGTTCGCGAGGCGCTCTTTTCATCATGCGAGCGTTCGATGAGTCAAGGCTTCCTT 2005
QY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
151359 TCTGGTCCAGAGGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 151418
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2006 CCGAGTTCGCGAGGCGCTCTTTTCATCATGCGAGCGTTCGATGAGTCAAGGCTTCCTT 2065
QY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
151419 ACATGGGCTAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 151478
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2066 CATTAAGTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2078
QY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
151479 CATTAAGTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 151538
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2079 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
151539 CATTAAGTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 151598
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```



```

CY      2079  -----AGGAGGTTCTGGCACTGACGAGGCGATTCTCTGTTGTTGGACTCTCTGAACGAGGA 2131
DE      151595  CTGAGAAAGGAAGTCTTGGCCACACGCTGCTGCTTCTGTTGGGCTCTCTAATGAAGA 151658
CY      2132  GATCAGGAGCCACCTGGAGAGAGTCTCTCTGAGAGTCTCTGCGACATGACATGAGATGGA 2191
DE      151659  GATGAGATGCTACCTCGAGAGGAATCTGCGCGGACCATCTTTCCTGCAAGSTTAAGGAGGA 151718
CY      2192  CTGTTGGCACTGGATCCAAAGCAAGGCTAGAGATGAAATGATGATGCTGATGATGAGAGGCTC 2251
DE      151719  AGGATTGGCGTGGATCCAAACCAAGCTCGAATGAGAGTCTTCTGCTGAGAGCATGCTC 151778
CY      2252  CTGCGAGTCTTTAGGCTGCTGATGAGATGCAAGAGATTAATTTATATAGGAGAGCT 2311
DE      151779  CCTGGAGCTCTGAGCTGTTGTAAGATGCAAGAGTAAGATCTATCCAGCAGGCTCT 151838
CY      2312  GAGCCATTTCAAGTATGCTGTTGAGCAATATGCTCTGAGATGAGATGATGCTCTC 2371
DE      151839  GAGCCCACTTTCAAGTGTGTTGTTGTCAGAAAGCTCTAAGCAAGATGAGAGCTGCTGCTG 151898
CY      2372  CTGTTGTTGTGAAGCGTTGCGAGAGAGGCTGAGGTTGCACTTGTATGGGCCCACTA 2431
DE      151899  CTGTTTGTGCAAGCTACTGAGAGGCAAGAAATGCTTACTTGTATGGAGTCCCTA 151958
CY      2432  CAGCGGAGGGGGAAGACCGCGAGAGTCTCTGATAGAGTGAACAGCTGT 2484
DE      151959  TATGACAGGGGCGAGAGATGCGCTGCGCAAAATCTTCAATGATGAGAGCACT 152011

```

Search completed: October 31, 2003, 09:15:42
 job time : 19581 secs

GenCore version 5.1.6
Copyright (c) 1993 2003 Coriagen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 30, 2003, 22:40:36 ; Search time 9464 Seconds
(without alignments)
12663.289 Million cell updates/sec

File: us-10-028-374-1

Perfect score: 4931

Sequence: 1 cccggcgccgaacatgggc.....aaaaaaaaaaaaaaaaaa 4931

Scoring table: IDENTITY_NJC

Gapop 10.0, Gapext 1.0

Searched: 2278192 seqs, 12152216056 residues

Total number of hits satisfying chosen parameters: 4556794

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: em_estba:*

2: em_estcm:*

3: em_estm:*

4: em_estm:*

5: em_estov:*

6: em_estpl:*

7: em_estto:*

8: em_estl:*

9: gb_estl:*

10: gb_est2:*

11: gb_hic:*

12: gb_estl:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: em_gss_hum:*

18: em_gss_inv:*

19: em_gss_pln:*

20: em_gss_vrt:*

21: em_gss_fun:*

22: em_gss_mar:*

23: em_gss_mus:*

24: em_gss_pro:*

25: em_gss_rtd:*

26: em_gss_phg:*

27: em_gss_vrt:*

28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result:	No.	Score	Query Match	Length	DB ID	Description
1	566.4	11.5	937	13	BQ903310	BQ903310 AGENCOURT
2	563.8	11.4	917	13	BQ903369	BQ903369 AGENCOURT
3	537.6	10.9	942	13	BQ907631	BQ907631 AGENCOURT
4	515.8	10.5	664	29	AG067258	Pan trogl

5	506.9	10.3	568	10	AW956628	EST:766898
6	377.4	7.7	439	28	AC014983	HS 5249.5
7	375.2	7.6	496	28	AC014983	HS 5249.5
8	362.6	7.4	396	28	AC094928	HS 3038.7A
9	338.6	6.9	449	28	AC044314	HS 5130.7A
10	324.4	6.6	788	12	BM043219	603619478
11	321.4	6.5	3584	11	BC036238	HS 603619478
12	311.4	6.3	764	12	BI011853	HS 603619478
13	301.2	6.1	660	29	AG169231	Pan trogl
14	291.2	6.0	571	1	AW074641	bab601.1
15	266.4	5.4	581	29	AG061279	Pan trogl
16	265.8	5.4	184	1	BC014171	HS 603619478
17	265.8	5.4	3527	11	BC040121	HS 603619478
18	265.8	5.4	3874	11	BC047576	HS 603619478
19	261.6	5.3	800	9	AV755512	AV755512
20	257.2	5.2	3146	11	BC036603	Homo sapi
21	255.4	5.2	633	13	EX471328	HS 603619478
22	255.4	5.2	822	10	BS245643	HS 603619478
23	255.4	5.2	1043	13	BM919569	AGENCOURT
24	254.6	5.2	1073	29	BZ601227	WHAC099TR
25	254.4	5.2	977	14	CD515864	AGENCOURT
26	254.2	5.2	393	10	BE139267	HS 603619478
27	253.2	5.1	685	28	AC033959	CIT-HSP 2
28	252.8	5.1	717	29	AG033865	Pan trogl
29	252.8	5.1	118	13	PJ536666	AGENCOURT
30	252.4	5.1	2044	11	BC037797	Homo sapi
31	251.4	5.1	576	2	HS067625	Homo sapi
32	251.4	5.1	357	13	BE139358	HS 603619478
33	251.4	5.1	421	13	BS054979	HS 603619478
34	250.8	5.1	530	28	BI4848	HS 603619478
35	250.6	5.1	638	28	AG043950	HS 603619478
36	250.4	5.1	670	29	AG035650	Pan trogl
37	250.4	5.1	883	13	BU181893	AGENCOURT
38	250.2	5.1	457	28	AG052735	CIT-HSP 1
39	250.2	5.1	553	28	BS11123	HS 603619478
40	250.2	5.1	1952	11	BC018692	HS 603619478
41	249.8	5.1	3554	11	BC032554	HS 603619478
42	249.6	5.1	584	28	AG033442	HS 603619478
43	249.6	5.1	666	29	AG084334	Pan trogl
44	249.6	5.1	669	29	AG182560	Pan trogl
45	249.4	5.1	653	29	AG159844	Pan trogl

ALIGNMENTS

RESULT 1
BQ903310
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

BQ903310
AGENCOURT AG067258 NIH_MGC_45 Homo sapiens cDNA clone IMAGE:6181082
5', mRNA sequence
BQ903310.1 GI:22692344
EST
Homo sapiens (human)
Eukaryote; Mammalia; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 bases 1 to 937
NIH_MGC http://mgc.nhlbi.nih.gov/
National Institutes of Health, Mar-Allen Gene Collection (MGC)
Unpublished
Contact: Robert Straussberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: DCTD/OTR
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: tlv2372 row: 1 column: 11
High quality sequence start: 5

Food...


```

Best Local Similarity 96.1%; Pred. No. 1e-02;
Matches 344; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 2073 TGTGAGAGAGTTCTGCTGCACTACAGAGCGCTTCCTGTTGGAGCTCTGAGAGAG 2132
DB 91 TGTGAGAGAGTTCTGCTGCACTACAGAGCGCTTCCTGTTGGAGCTCTGAGAGAG 150
QY 2133 ACCAGAGAGAGTTCTGCTGCACTACAGAGCGCTTCCTGTTGGAGCTCTGAGAGAG 2192
DB 151 ACCAGAGAGAGTTCTGCTGCACTACAGAGCGCTTCCTGTTGGAGCTCTGAGAGAG 210
QY 2193 CTGTTGCACTGATCAAGCAAGAGTTCTGCTGCACTACAGAGCGCTTCCTGTTGGAGCTCTGAGAGAG 2252
DB 211 CTGTTGCACTGATCAAGCAAGAGTTCTGCTGCACTACAGAGCGCTTCCTGTTGGAGCTCTGAGAGAG 270
QY 2253 TTTGAGAGTTCTGCTGCACTACAGAGCGCTTCCTGTTGGAGCTCTGAGAGAG 2312
DB 271 TTTGAGAGTTCTGCTGCACTACAGAGCGCTTCCTGTTGGAGCTCTGAGAGAG 330
QY 2313 ACCAGAGAGAGTTCTGCTGCACTACAGAGCGCTTCCTGTTGGAGCTCTGAGAGAG 2372
DB 331 ACCAGAGAGAGTTCTGCTGCACTACAGAGCGCTTCCTGTTGGAGCTCTGAGAGAG 390
QY 2373 TGTGTTGCTGCACTGATCAAGCAAGAGTTCTGCTGCACTACAGAGCGCTTCCTGTTGGAGCTCTGAGAGAG 2430
DB 391 TGTGTTGCTGCACTGATCAAGCAAGAGTTCTGCTGCACTACAGAGCGCTTCCTGTTGGAGCTCTGAGAGAG 448

RESULT 10
LOCUS NM043219
DEFINITION 603619428F1 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:543283 5'.
ACCESSION NM043219
VERSION 1
KEYWORDS EST
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE NIH-MGC http://mgc.nci.nih.gov/
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgabbs@mail.nih.gov
Tissue procurement: DCTD/NIH
cDNA library preparation: NIH High/Through Laboratory
cDNA library Arrayed by: The I.M.A.G.E. Consortium (ILM)
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium (ILM) at:
http://image.llnl.gov
Plate: LLC1907 row: 9 column: 04
High quality sequence stop: 653.
Location/Qualifiers
1..788
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:543283"
/tissue_type="carcinoma, cell line"
/lab_host="DHIOH (Phage T4 strain)"
/clone_lib="NIH_MGC_40"
/notes="Organ: Pituitary, Vector: pMTS7; Site 1: XhoI;
Site 2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GCCACGAGG. Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH MGC Library."
BASE COUNT 177 a 211 c 243 g 157 t

```

```

ORIGIN
Query Match 6.6%; Score 324.4; DB 12; Length 788;
Best Local Similarity 85.5%; Pred. No. 1.2e-02;
Matches 584; Conservative 0; Mismatches 56; Indels 30; Gaps 12;

QY 2073 TGTGAGAGAGTTCTGCTGCACTACAGAGCGCTTCCTGTTGGAGCTCTGAGAGAG 2132
DB 76 TGTGAGAGAGTTCTGCTGCACTACAGAGCGCTTCCTGTTGGAGCTCTGAGAGAG 135
QY 2133 ACCAGAGAGAGTTCTGCTGCACTACAGAGCGCTTCCTGTTGGAGCTCTGAGAGAG 2192
DB 136 ACCAGAGAGAGTTCTGCTGCACTACAGAGCGCTTCCTGTTGGAGCTCTGAGAGAG 195
QY 2193 CTGTTGCACTGATCAAGCAAGAGTTCTGCTGCACTACAGAGCGCTTCCTGTTGGAGCTCTGAGAGAG 2250
DB 196 CTGTTGCACTGATCAAGCAAGAGTTCTGCTGCACTACAGAGCGCTTCCTGTTGGAGCTCTGAGAGAG 255
QY 2251 CTTTGGAGTTCTGCTGCACTACAGAGCGCTTCCTGTTGGAGCTCTGAGAGAG 2309
DB 256 CTTTGGAGTTCTGCTGCACTACAGAGCGCTTCCTGTTGGAGCTCTGAGAGAG 315
QY 2310 CTGAGCGAG-TTCAGAGTTCTGCTGCACTACAGAGCGCTTCCTGTTGGAGCTCTGAGAGAG 2368
DB 316 CTGAGCGAGTTCTGCTGCACTACAGAGCGCTTCCTGTTGGAGCTCTGAGAGAG 375
QY 2369 CTGCTGCTGTTCTGCTGCACTACAGAGCGCTTCCTGTTGGAGCTCTGAGAGAG 2427
DB 376 CTGCTGCTGTTCTGCTGCACTACAGAGCGCTTCCTGTTGGAGCTCTGAGAGAG 435
QY 2428 CTTTGA-TTCAGAGTTCTGCTGCACTACAGAGCGCTTCCTGTTGGAGCTCTGAGAGAG 2483
DB 436 CTTTGA-TTCAGAGTTCTGCTGCACTACAGAGCGCTTCCTGTTGGAGCTCTGAGAGAG 495
QY 2484 TTGCTGTTCTGCTGCACTACAGAGCGCTTCCTGTTGGAGCTCTGAGAGAG 2535
DB 496 TTGCTGTTCTGCTGCACTACAGAGCGCTTCCTGTTGGAGCTCTGAGAGAG 555
QY 2536 TG-GCAGAGCGCTTCCTGCACTACAGAGCGCTTCCTGTTGGAGCTCTGAGAGAG 2584
DB 556 TG-GCAGAGCGCTTCCTGCACTACAGAGCGCTTCCTGTTGGAGCTCTGAGAGAG 615
QY 2585 CCGAATGCTGCTGCACTACAGAGCGCTTCCTGTTGGAGCTCTGAGAGAG 2636
DB 616 GGAATGCTGCTGCACTACAGAGCGCTTCCTGTTGGAGCTCTGAGAGAG 669

RESULT 11
LOCUS BC036238
DEFINITION Homo sapiens, clone IMAGE:5297327, mRNA.
ACCESSION BC036238
VERSION BC036238.1
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE Direct Submission
JOURNAL Strausberg, R.
COMMENT Submitted (31-JUL-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 1A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgabbs@mail.nih.gov
Tissue Procurement: Niklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael C. Brownstein (NHGRI) & Shiraki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILM)
cDNA Sequencing by: Sequencing Group at the Stanford Human Genome

```


Clone distribution: VSC clone distribution information can be found through the I.M.A.G.E. Consortium/INM at: <http://image.llnl.gov>

Plate: L2AM1534 row: C column: 10
High quality sequence start: 23
High quality sequence stop: 764

FEATURES

Location/Qualifiers
1..764
/organism="Homo sapiens"
/mol_type="trna"
/db_xref="taxon:9606"
/clone="IMAGE:5212737"
/issue_type="leukocyte"
/lab_host="DH1CB"
/clone_lib="NIH MSC LIB"
/note="Vector: PMV SPART6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA source: leukocytes from anonymous pool of non-activated adult donors. Library is oligo dT primed and directionally cloned. EcoRV site is destroyed upon cloning. Average insert size 1.7 kb, insert size range 1.2-3.3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 527. Note: this is a NIH MSC Library."
173 a 224 c 234 g 133 c

BASE COUNT

ORIGIN 173 a 224 c 234 g 133 c
Query Match 6.3%; Score 116; DB 12; Length 764;
Best Local Similarity 99.4%; Pred. No. 2,20+02;
Matches 132; Conservative 2; Mismatches 0; Indels 2; Gaps 2;

QY 940 GATACCCAGCTGTGGCGCGCTTACTTCTTGGACAGATGACATGCTTGGAGAGC 999
DB 411 GATACCCAGCTGTGGCGCGCTTACTTCTTGGACAGATGACATGCTTGGAGAGC 999
QY 1000 TCTCTTCTCACTCAAGAAAGATGCTTCTTGGACAGATGACATGCTTGGAGAGC 1059
DB 431 TCTCTTCTCACTCAAGAAAGATGCTTCTTGGACAGATGACATGCTTGGAGAGC 1059
QY 1260 TTCTGGTCTATGAG 1118
DB 551 TTCTGGTCTATGAG 1118
QY 1119 GTACACCGGCTCTGT 1178
DB 611 GTACACCGGCTCTGT 1178
QY 1179 TCTGGACACAGCGCGGAG 1237
DB 571 TCTGGACACAGCGCGGAG 1237
QY 1238 TAGAGACCTCTTTGAGCCAG 1271
DB 731 TAGAGACCTCTTTGAGCCAG 1271

RESULT 13

AG169231/c
DEFINITION Pan troglodytes DNA, clone: RP43-037N12.7, genomic survey
sequence. GSS 69-JAN-2002

ACCESSION

AG169231
VERSION AG169231.1 GI:16698909
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)

ORGANISM

Eukaryota; Metazoa; Chordata; (Mammalia; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Pan; Hominidae; Homo; Homo sapiens)

REFERENCE

1. Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T., Torok, Y., Watanabe, H. and Sakaki, Y. BAC end sequences of Library RP43-037N12.7. J. Biol. Chem. 278:10000-10004 (2003).
Unpublished

REFERENCE

2. (bases 1 to 660)
Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T., Torok, Y., Watanabe, H. and Sakaki, Y. BAC end sequences of Library RP43-037N12.7. J. Biol. Chem. 278:10000-10004 (2003).
Unpublished

CONVENT

Clones are derived from the chimpanzee BAC library RP43-037N12.7 and was generated during the R&D process and may have higher chance of clone tracking errors.

FEATURES

source
Location/Qualifiers
1..660
/organism="pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/clone="RP43-037N12.7"
/sex="male"
/cell_type="lymphocytes"
/clone_lib="RP43-037N12.7"
151 a 170 c 165 g 173 t 1 others

Query Match

Best Local Similarity 79.6%; Pred. No. 3,50+02;
Matches 412; Conservative 0; Mismatches 103; Indels 9; Gaps 4;

QY 3570 TTTTCTTTTGGAG 3629
DB 599 TTTTCTTTTGGAG 541
QY 3630 AA TTTTCTTTTGGAG 3683
DB 540 AATTTCTTTTAAATTTTGGAG 481
QY 3684 CAGTCTATGAG 3743
DB 480 CAGTCTATGAG 421
QY 3744 TGAAGCAGTGTGAG 3823
DB 420 TGAAGCAGTGTGAG 361
QY 3904 TGGGTGAAGTCTTAAATTTGGAG 3963
DB 360 TGGGTGAAGTCTTAAATTTGGAG 301
QY 3864 GGGATTTGAG 3923
DB 300 GGGATTTGAG 243
QY 3924 AAGTAGCAG 3983
DB 242 GATTTAGCAG 183
QY 3984 AGAATCCCTTTGAAATCCCGGAG 4043
DB 182 AGAATCCCTTTGAAATCCCGGAG 123
QY 4044 CAGTCTGAG 4087
DB 122 CAGTCTGAG 79

RESULT 14

AM673661/c

IP 469 AAAGGGGGGGGGGAAACCCGGAGGGGTTTCAGAAAGAGTCCACCCCTGTGGGC 526

Search completed: October 31, 2003, 06:57:36
Query time : 9474 secs

GenCore Version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

CM protein - protein search, using sw model

Run on: October 31, 2003, 01:57:44 : Search time 42 seconds
without alignment
263.125 Million cell updates/sec

Title: US-10-028-374-2

Perfect score: 237

Sequence: 1 MNQATECSMQLFSCWPE.....GVKILVQLRHPEKCKQLR 449

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 1137863 seqs, 158724573 residues

Total number of hits satisfying chosen parameters: 1:07953

Minimum DB seq length: 3

Maximum LP seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: A Geneset 10Jun01*

1	/SIDSI/gcgdata/geneseq/geneseqp-emb/AA1960.DAT*
2	/SIDSI/gcgdata/geneseq/geneseqp-emb/AA1961.DAT*
3	/SIDSI/gcgdata/geneseq/geneseqp-emb/AA1962.DAT*
4	/SIDSI/gcgdata/geneseq/geneseqp-emb/AA1963.DAT*
5	/SIDSI/gcgdata/geneseq/geneseqp-emb/AA1964.DAT*
6	/SIDSI/gcgdata/geneseq/geneseqp-emb/AA1965.DAT*
7	/SIDSI/gcgdata/geneseq/geneseqp-emb/AA1966.DAT*
8	/SIDSI/gcgdata/geneseq/geneseqp-emb/AA1967.DAT*
9	/SIDSI/gcgdata/geneseq/geneseqp-emb/AA1968.DAT*
10	/SIDSI/gcgdata/geneseq/geneseqp-emb/AA1969.DAT*
11	/SIDSI/gcgdata/geneseq/geneseqp-emb/AA1970.DAT*
12	/SIDSI/gcgdata/geneseq/geneseqp-emb/AA1971.DAT*
13	/SIDSI/gcgdata/geneseq/geneseqp-emb/AA1972.DAT*
14	/SIDSI/gcgdata/geneseq/geneseqp-emb/AA1973.DAT*
15	/SIDSI/gcgdata/geneseq/geneseqp-emb/AA1974.DAT*
16	/SIDSI/gcgdata/geneseq/geneseqp-emb/AA1975.DAT*
17	/SIDSI/gcgdata/geneseq/geneseqp-emb/AA1976.DAT*
18	/SIDSI/gcgdata/geneseq/geneseqp-emb/AA1977.DAT*
19	/SIDSI/gcgdata/geneseq/geneseqp-emb/AA1978.DAT*
20	/SIDSI/gcgdata/geneseq/geneseqp-emb/AA1979.DAT*
21	/SIDSI/gcgdata/geneseq/geneseqp-emb/AA1980.DAT*
22	/SIDSI/gcgdata/geneseq/geneseqp-emb/AA1981.DAT*
23	/SIDSI/gcgdata/geneseq/geneseqp-emb/AA1982.DAT*
24	/SIDSI/gcgdata/geneseq/geneseqp-emb/AA1983.DAT*

Proj. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARY

Result No.	Score	Match	Length	DB ID	Description
1	2377	100.0	449	23	ABR77910 Amino acid sequence
2	2344.5	98.6	1561	23	AAAC15690 Human PYRIN-3 prot
3	2334	98.2	1099	23	AAAC17857 Pyrin domain conta
4	1157.5	48.7	1034	22	AAAC27514 Human PYRIN-1 prot
5	1157.5	48.7	1034	24	AAAC28023 Huma PYRIN-1 prote
6	1039.5	43.7	565	23	ABG78474 Human nucleic acid
7	853	35.9	168	22	ABG04570 Novel human diago
8	816	34.3	920	23	ASP54254 Human MODT-13 prot
9	810	34.1	518	23	AAAC17871 Pyrin domain conta

10	759	31.9	1397	22	AA72670	Human NB ARC and C
11	759	31.9	1424	22	AA72674	Human NAC gamma of
12	759	31.9	1429	22	AA862571	Human CARD 2 polyp
13	759	31.9	1429	23	ABG7969	Human leucine rich
14	759	31.9	1429	23	ABG78472	Leucine-rich repea
15	759	31.9	1429	23	ABG77916	Human leucine-rich
16	759	31.9	1429	24	ABG71631	Human caspase recr
17	759	31.9	1429	24	ABG71631	Human caspase recr
18	759	31.9	1429	24	ABG71631	Human NB-ARC and C
19	759	31.9	1429	24	AA72674	Human NAC beta 1sc
20	759	31.9	1429	22	AA72673	Human G-protein: cc
21	759	31.9	1429	22	AA72669	Human NB-ARC and C
22	759	31.9	1429	22	AA72712	Human NAC beta 1sc
23	759	31.9	1429	23	ABG78455	Human caspase recr
24	759	31.9	1429	23	AAAC17855	Pyrin domain conta
25	759	31.9	1429	23	AAAC17858	Pyrin domain conta
26	714.5	30.1	484	22	AAAC1096	Gene 28 Human secr
27	714.5	30.1	629	24	ABG01160	Human gene 214 enc
28	714.5	30.1	630	22	AAAC1087	Human secreted pro
29	737.5	29.8	996	23	AAAC15592	Human PYRIN-11 pro
30	694.5	29.2	514	20	AA739778	CRDAX101 protein s
31	694.5	29.2	514	23	ABG7969	Human leucine rich
32	694.5	29.2	514	23	ABG78471	Leucine rich repea
33	694.5	29.2	514	23	ABG77915	Arginylserinevasop
34	684	28.8	676	23	ABG65675	Human NOV18a prote
35	675	28.4	891	22	AAAC04546	Human G-protein: cc
36	652	27.9	159	21	AAAC12357	Human ORF2121
37	652.5	27.5	858	23	AAAC3328	Human nucleotide b
38	652.5	27.5	919	23	ABP43483	Human secreted pro
39	652.5	27.5	913	24	ABP19319	NOVX related prote
40	652.5	27.5	933	23	AAAC17870	Pyrin domain conta
41	652.5	27.5	934	23	AAAC15593	Human PYRIN-3 prot
42	649	27.3	1551	22	AAAC2022	Human activated T-
43	649	27.3	1551	22	AAAC17860	Pyrin domain conta
44	641	27.0	625	23	ABG78454	Human leucine rich
45	627.5	26.4	2312	23	AAAC17859	Pyrin domain conta

ALIGNMENTS

RESULT 1

ABR77910
ID ABR77910 standard: Protein: 449 AA.

AC ABR77910

XX ABR77910

XX ABR77910

XX ABR77910

XX ABR77910

DE Amino acid sequence of human leucine-rich repeat protein HLRPM1.

XX Human leucine-rich repeat; HLRPM1; proliferative disorder;

XX immune condition; apoptosis; signal transduction; autoimmune disease;

XX haematopoietic cell disease; graft versus-host disease; allergy; asthma;

XX cardiovascular disorder; neurological disease; pheromone;

XX pulmonary disease; chronic obstructive pulmonary disease;

XX allergic rhinitis; bronchial hyperresponsiveness; reproductive disease;

XX haematopoietic disease; platelet disorder; Bernard-Soulier syndrome;

XX inflammatory disorder; systemic lupus erythematosus;

XX cardiovascular disease; cancer.

OS Homo sapiens.

XX Homo sapiens.

PH Key location/Qualifiers

FT Domain 141-159

XX Domain 141-159

XX Domain 141-159

XX Domain 141-159

XX Domain 141-159

XX Domain 141-159

XX Domain 141-159

XX Domain 141-159

XX Domain 141-159

XX Domain 141-159

XX Domain 141-159

XX Domain 141-159

XX Domain 141-159

XX Domain 141-159

XX Domain 141-159

22-DEC-2003; 2000US-257773P.
(PRIM) BRISTOL-MYERS SQUIBB CO.
Foot: C. Ramaratnan C. Mintier G;
WPI: 2002-56656/60.
N-PSDB; AAL59333.
New HLRRPVI nucleic acids for preventing, treating or ameliorating e.g. proliferative disorders, immune conditions, a disorder related to aberrant apoptosis modulation or developmental disorders
Claim 20; Fig 1A-B; 31pp; English.
The present sequence represents a human leucine-rich repeat containing protein, designated HLRRPVI. HLRRPVI polypeptides and polynucleotides are useful for preventing, treating or ameliorating a medical condition such as a proliferative disorder, immune condition, or a disorder related to aberrant apoptosis modulation, either directly or indirectly, and in modulating signal transduction activity in various cells, tissue and organisms. They are also useful for treating, preventing, or diagnosing diseases of haematopoietic cells, autoimmune disease, graft-versus-host disease, allergic conditions (e.g. asthma), cardiovascular disorders, and neurological diseases, and for increasing the organism's ability to synthesize and/or release hormones. The polypeptide may also be used in treating, preventing or ameliorating pulmonary disease (e.g. chronic obstructive pulmonary disease, allergic rhinitis, or bronchial hyperresponsiveness), reproductive disease, haematopoietic disease, platelet disorders (e.g. Bernard Soulier syndrome), non-infectious disorders (e.g. innate immunity to bacterial pathogens, or adaptive immune response), immune and inflammatory disorders (e.g. systemic lupus erythematosus), cardiovascular diseases and cancers. HLRRPVI nucleic acids may further be used in chromosome identification or mapping, as a chromosome marker, as molecular weight markers, as diagnostic probes, in gene therapy, in raising anti-DNA antibodies, or as antigens for eliciting immune responses.

DB 42: LYRNALGSPGVKLLCQQLRHHPNCKLNL 449

RESULT 2
AAC15590
ID AAC15590 standard. Protein: 1061 AA.
XX
XX AAC15590:
XX
XX 31-OCT-2002 (first entry)
XX
XX Human PYRN-8 protein #2.
XX
XX Human; Gene therapy; PYRN; stress-related response; apoptotic response; inflammatory response; inflammatory disorder; immune system disorder; Crohn's disease; multiple sclerosis; cancer; leukaemia;
XX
XX autoimmune disorder; arthritis; neurological disease;
XX
XX Alzheimer's disease; Parkinson's disease; chromosome mapping;
XX
XX tissue typing; forensic biology; predictive medicine; pharmacogenomics;
XX
XX transcription profiling; PYRN-8.
XX
XX Homo sapiens.
XX
XX WC200201049-AA.
XX
XX 08-AUG-2002.
XX
XX 31-JAN-2003; 2000WO-US29637.
XX
XX 31-JAN-2001; 2001US-265231P.
XX
XX 10-SEP-2001; 2001US-318645P.
XX
XX (MILL) MILLERKUN PHARM INC.
XX
XX (AMRP) WYETH.
XX
XX Bertin J, Wang W, Blatcher M,
XX
XX WPI: 2002-627478767.
XX
XX N-PSDB; AAL44363.
XX
XX New PYRN polypeptides and nucleic acids useful for modulating and
XX
XX diagnosing stress-related, apoptotic and inflammatory responses, or for
XX
XX treating inflammatory and immune system disorders, cancers, or
XX
XX neurological diseases.
XX
XX Claim 8; Fig 5; 16pp; English.
XX
XX The invention comprises the amino acid and coding sequences of human
XX
XX PYRN proteins. The PYRN protein and DNA sequences of the invention are
XX
XX useful for modulating and diagnosing stress-related, apoptotic and
XX
XX inflammatory responses. The PYRN protein and DNA sequences are useful
XX
XX for treating inflammatory disorders and immune system disorders (e.g.
XX
XX Crohn's disease, reactive arthritis, multiple sclerosis, contact
XX
XX dermatitis, psoriasis, graft rejection, allergies, viral infections and
XX
XX bacterial infections); cancer (e.g. leukaemia); autoimmune disorders
XX
XX (e.g. systemic lupus erythematosus and arthritis); and neurological
XX
XX diseases (e.g. Alzheimer's disease and Parkinson's disease). The PYRN
XX
XX protein and DNA sequences may also be used in screening assays, detection
XX
XX assays (e.g. chromosomal mapping, tissue typing or forensic biology),
XX
XX predictive medicine (e.g. diagnostic assays, clinical trials and
XX
XX pharmacogenomics) and transcription profiling. The present amino acid
XX
XX sequence represents a human PYRN-8 protein.

Query Match 100.0%; Score 2344.5; DB 23; Length 1061;
Best Local Similarity 90.9%; Pred. No. 2.8e-232;
Matches 449; Conservative 3; Mismatches 0; Indels 45; Gaps 20

QY 1 MNQATGCSMCQJFSCWPFSPAPLQELIRVPERLFTINGFCEIAKSPEDFGGPKCLCW 60
DB 1 MNQATGCSMCQJFSCWPFSPAPLQELIRVPERLFTINGFCEIAKSPEDFGGPKCLCW 60
QY 61 EERKPTIELUNSLIRKLLPEGLSITTTPTALEKHLNFIIEHPHVEILGFEAEKKEVF 120
DB 61 EERKPTIELUNSLIRKLLPEGLSITTTPTALEKHLNFIIEHPHVEILGFEAEKKEVF 120
QY 121 YKVFHRAEGAGQVNVVQNEPFTFWCVPLVWVVTLCQQLGGGLRQTSPTITAV 180
DB 121 YKVFHRAEGAGQVNVVQNEPFTFWCVPLVWVVTLCQQLGGGLRQTSPTITAV 180
QY 181 YMLYLLSLQPKPKAPLQPPKPGELCSIAATGLANQFIIIEEQELRHKHLGGEDYSAF 240
DB 181 YMLYLLSLQPKPKAPLQPPKPGELCSIAATGLANQFIIIEEQELRHKHLGGEDYSAF 240
QY 241 LNNKIFOKQINCESEFALTSRFLFQJLKHETSHLEKSLWVVSPIIKVLLQWIOSKA 300
DB 241 LNNKIFOKQINCESEFALTSRFLFQJLKHETSHLEKSLWVVSPIIKVLLQWIOSKA 300
QY 301 QSGDSTLQGSLEFFSCVLEIEEERFQALNHFQVIVNSIASKMHWVSFCLKRCRS 360
DB 301 QSGDSTLQGSLEFFSCVLEIEEERFQALNHFQVIVNSIASKMHWVSFCLKRCRS 360
QY 361 AQVHLHYGATYSADGEDRARCAGATLLVQLPFTVLLQAYSEHAAALCNPNLEIUS 420
DB 361 AQVHLHYGATYSADGEDRARCAGATLLVQLPFTVLLQAYSEHAAALCNPNLEIUS 420
QY 421 LYRNALGSPGVKLLCQQLRHHPNCKLNL 449

[illegible]

FN WC200272630-A2.

QY 258 ALTSRFLGGLNEETSRSHLEKSLQKWSPH:KMDLLQWIOQKASDGSSTLQGSLEFFSC 317
 DB 61 ALTSRFLGGLNEETSRSHLEKSLQKWSPH:KMDLLQWIOQKASDGSSTLQGSLEFFSC 120
 QY 318 LYEIOEEEFIOQALSHFQVIVYSNIASKVHMVSSFCFKPFEACQVILHYGATYSAGDC 377
 DB 121 LYEIOEEEFIOQALSHFQVIVYSNIASKVHMVSSFCFKPFEACQVILHYGATYSAGDC 180
 QY 378 RARCSAGANTLLVQVLPERTVLDAYSEHLAAALCTNPNIIELSLYR 437
 DB 181 RARCSAGANTLLVQVLPERTVLDAYSEHLAAALCTNPNIIELSLYR 240
 QY 438 LRHPNCKLQNP 449
 DB 241 LRHPNCKLQNP 252
 RESULT 7
 AR30457C
 ID AB304570 standard; Protein: 168 AA.
 XX AC AR30457C;
 XX DT 13-FEB-2002 (first entry;
 XX XX Move: human diagnostic protein #4561.
 XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 XX KW food supplement; medical imaging; diagnostics; genetic disorder.
 XX XX Homo sapiens.
 XX XX W:200175067-A2.
 XX XX 11-OCT-2001.
 XX XX 30 JAN 2001; 2001WC-USC4631.
 XX XX 11-MAR 2000; 2000JS-0540217.
 XX XX 23 AUG-2000; 2000JS-0649167.
 XX XX (HYPSE-) HYSEQ INC.
 XX XX Drmanac RT, Liu C, Tang YT;
 XX XX WPI: 2001-639362/73.
 XX XX N-PSDB; AA568757.
 XX PT New isolated polynucleotide and encoded polypeptides, useful in
 XX PT diagnostics, forensics, gene mapping, identification of mutations
 XX PT responsible for genetic disorders or other traits and to assess
 XX PT biodiversity.
 XX PS Claim 20; SEQ ID No 34923; 103pp; English.
 XX CC The invention relates to isolated polynucleotides (i) and
 XX CC polypeptide (ii) sequences. (i) is useful as hybridisation probes,
 XX CC polymerase chain reaction (PCR) primers, oligonucleotides, and for chromosome
 XX CC and gene mapping, and in recombinant production of (ii). The
 XX CC polynucleotides are also used in diagnostics as expressed sequence tags
 XX CC for identifying expressed genes. (ii) is useful in gene therapy techniques
 XX CC to restore normal activity of (i) or to treat disease states involving
 XX CC (iii). (iii) is useful for generating antibodies against it, detecting or
 XX CC quantitating a polypeptide in tissue, as molecular weight markers and as
 XX CC a food supplement. (ii) and its binding partners are useful in medical
 XX CC imaging of sites expressing (ii). (i) and (ii) are useful for treating
 XX CC disorders involving aberrant protein expression or biological activity.
 XX CC The polypeptide and polynucleotide sequences have applications in
 XX CC diagnostics, forensics, gene mapping, identification of mutations
 XX CC responsible for genetic disorders or other traits to assess biodiversity
 XX CC and to produce other types of data and products dependent on DNA and
 XX CC amino acid sequences. ABG0010-ABG30377 represent novel human
 XX CC diagnostic amino acid sequences of the invention.

CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX SQ Sequence 168 AA;
 Query Match 35.9%; Score 853; DB 22; Length 168;
 Best Local Similarity 59.4%; Pred. No. 1.7e-79;
 Matches 167; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 256 FLALTSRFLGGLNEETSRSHLEKSLQKWSPH:KMDLLQWIOQKASDGSSTLQGSLEFF 315
 DB 1 FLALTSRFLGGLNEETSRSHLEKSLQKWSPH:KMDLLQWIOQKASDGSSTLQGSLEFF 60
 QY 316 SCUYEQEEEFIOQALSHFQVIVYSNIASKVHMVSSFCFKPFEACQVILHYGATYSAGD 375
 DB 61 SCUYEQEEEFIOQALSHFQVIVYSNIASKVHMVSSFCFKPFEACQVILHYGATYSAGD 120
 QY 376 EDRARCSAGANTLLVQVLPERTVLDAYSEHLAAALCTNPNIIELSLYR 423
 DB 121 EDRARCSAGANTLLVQVLPERTVLDAYSEHLAAALCTNPNIIELSLYR 168
 RESULT 8
 ABP53254
 ID ABP53254 standard; Protein: 920 AA
 XX AC ABP53254;
 XX DT 13-NOV-2002 (first entry;
 XX XX Human MDDT-13 protein SEQ ID NO:13.
 XX KW Human; MDDT; molecule for disease detection and treatment; anti-HIV;
 KW antiarteriosclerotic; cytoskeletal; neuroprotective; antiparkinsonian;
 KW hepatotropic; laxative; cerebroprotective; antiinflammatory; virucide;
 KW antibacterial; fungicide; protozoicide; gene therapy; arteriosclerosis;
 KW cirrhosis; cancer; stroke; Alzheimer's disease; Parkinson's disease;
 KW Crohn's disease; constipation; acquired immunodeficiency syndrome; AIDS;
 XX XX infection.
 XX XX Homo sapiens.
 XX XX WC200264792-A2.
 XX XX 22-AUG-2002
 XX XX 04-JAN-2002; 2002WC-US00254.
 XX XX 05-JAN-2001; 2001US-260168P.
 XX XX 19-JAN-2001; 2001US-262366P.
 XX XX 19-JAN-2001; 2001US-262857P.
 XX XX (INCYTE) INCYTE GENOMICS INC.
 XX XX Baughn MR, Warren BA, Ronche JJ, Xu Y, Wallia NK, Ramkumar J;
 XX XX Rao MG, Lu Y, Yue H, Thangaveeru K, Tang Y, Ding L, Borowsky M;
 XX XX Hafalila ACA, Lu DAW, Azizai Y, Tran B, Nguyen DB, Burford N;
 XX XX Ison GH, Gururajan R, Gandhi AR, Elliott VS, Tran UK;
 XX XX WPI: 2002-674890/72.
 XX XX N-PSDB; ABQ75811.
 XX PT New purified polypeptides, molecules for disease detection and
 XX PT treatment (MDDT), useful for diagnosing or treating a disease or
 XX PT condition associated with decreased expression or overexpression of
 XX PT functional MDDT e.g. stroke.
 XX PS Claim 1; Page 128-130; 162pp; English.
 XX CC ABQ75799 to ABQ75814 encode the human molecules for disease detection
 XX CC and treatment: 1 to 26 (MDDT-1 to MDDT-26) given in ABP53242 to ABP53267.
 XX CC MDDT's have anti-HIV, antiarteriosclerotic, cytostatic, neuroprotective,

CC antiParkinsonian, hepatotropic, laxative, neuroprotective, virucide,
 CC antiinflammatory, antibacterial, fungicide and protease activities,
 CC and can be used in gene therapy. MDP sequences can be used in the
 CC diagnosis and treatment of a disease or condition associated with
 CC decreased expression or overexpression of functional MDP e.g. cirrhosis,
 CC arteriosclerosis, cancer, stroke, Alzheimer's disease, Parkinson's
 CC disease, Crohn's disease, constipation, acquired immunodeficiency
 CC syndrome (AIDS), or bacterial, viral, fungal or protozoal infections.
 XX
 XX Sequence 920 AA;
 Query Match 34.3%; Score 816; DB 23; Length 920;
 Best Local Similarity 38.5%; Pred. No. 1 to 74;
 Matches 192; Conservative 76; Mismatches 159; Indels 72; Gaps 10;
 QY 7 ESMQQLFSCWPEPSAPQELIRVPERLFTIDGDELKPSFHDQGGPWCCLWKEKRP 66
 Db 255 ERSFAGLAKMSTFSPHEENYQSSSLPIHSFDELNFAFEPEFALCEDWTQEPV 314
 QY 67 ELLNSLRKALPELSLITTPFALEKIHLELHPHVEILGSEAEKRYFYKPHN 126
 Db 315 SFVSSLLKRWLPASLLVTRITTSKRLQQLKHHYVLELXSDAREEYVQPFED 374
 QY 127 AQCAGYVYVNDNEPLFTMCWPLVWVATCLQQGLKSLLELQTSKTTAVVNL 186
 Db 375 KRWKVPSSKSNELFSCVPLVWAACTLQKQWKGSDVTTCCTTAUFTCVIS 434
 QY 157 LQMCP-KPGAPLQPPNG---KGLSLAAHGMNKLLEFECQDPKHGLGEGDVSALN 242
 Db 435 SLFTPWGSPSL---PNAQLERLQVARGIWTYVYFRENRLRLGTCSSVSSND 491
 QY 243 MNIFQDI--NCE:-----RSLA-TSR-- 262
 Db 492 SNTCKDAEYENYVPTHELVETFAANFYMLKSNWAGNQCQPFEDKLSLQSTSYKD 551
 QY 263 -----ELFGLAETSSLEKSLQWVSHLKMGLQNLQSKAQDGSFLQGSLEFF 612
 Db 552 PLITQWKCFLGLEDNRYKQLETFNCKNSKIKSLQNEVLNGLNDSYSPSLGFL 612
 QY 315 FSLVETQREETIQALSHFQVIVYNTASKYHVVSPFLKRCDSAGVHNGATVSAD 374
 Db 612 FPLCYETOKAFISGAMRCFPAVNIPTKSHLVSSFLKRCLEPTLR---SVTV 667
 QY 375 GEDRACSGAGTLLVQLEPT---VLDKAYNHIAAATNPNLKLSNLYKHALGSR 429
 Db 665 FEKK---LKTSLPTNMLEYGLSLQCLLSALIQKSLKNNMLTQIGYE 719
 QY 430 GVKLLQGG-RHFNCK-QNL 448
 Db 720 GIVKLYKVKSPCKLOVL 738
 RESULT 9
 AAQ17871
 ID AAQ17871 standard; Protein; 518 AA.
 XX
 AC AAQ17871;
 XX
 XX
 XX 20-AUG-2002 (first entry);
 XX Pyrin domain containing protein NALP12/PY20 m.
 XX
 XX Pyrin domain; PYD domain; antiinflammatory; antiParkinsonian;
 XX antiarteriosclerotic; antiparasitic; antibacterial; virucide;
 XX neuroprotective; antirheumatic; antineuritic; antiasthmatic;
 XX nephrotropic; osteopathic; neurotropic; intracellular signal transduction;
 XX inflammation; Alzheimer's disease; infectious psoriasis; asthma;
 XX arteriosclerosis; multiple sclerosis; rheumatoid arthritis; sarcoidosis;
 XX Sarcoid arthritis; glomerulonephritis.
 XX
 XX Unidentified.
 XX
 XX WD0202040669-A2.

XX 23-MAY-2002.
 XX 30-OCT-2001; 2001WC HP12545.
 XX
 PR 15-NOV-2000; 2000DE-1556697.
 PR 30-NOV-2000; 2000DE-1059595.
 XX
 PA (APCT-); APOTECH RES & DEV LTD.
 XX
 PI Tschopp C, Martenson F;
 XX WP; 2002-427051/45;
 DR N-PSDB; AAL47143.
 XX
 PT New DNA encoding protein with pyrin domain, useful for treating
 PT diseases involving impaired signal transduction, particularly
 PT inflammation, also proteins and antibodies.
 XX
 PS Claim 5; Fig 1; 116pp; Orman.
 XX
 CC The present invention relates the DNA and their encoded proteins, where
 CC the proteins contain at least one PYD (pyrin) domain. These can be used
 CC to treat diseases associated with impaired intracellular signal
 CC transduction, particularly inflammation such as psoriasis,
 CC arteriosclerosis, bacterial or viral infections (particularly meningitis
 CC and pneumonia), multiple sclerosis, rheumatoid arthritis, asthma,
 CC sarcoidosis, glomerulonephritis and osteoarthritis, and also Alzheimer's
 CC and Parkinson's diseases. The present sequence is a protein of the
 CC invention.
 XX
 SQ Sequence 518 AA;
 Query Match 34.1%; Score 810; DB 23; Length 518;
 Best Local Similarity 60.6%; Pred. No. 2 to 74;
 Matches 152; Conservative 42; Mismatches 57; Indels 2; Gaps 0;
 QY 6 TEESMQLLPSWPEPSAPQELIRVPERLFTIDGDELKPSFHDQGGPWCCLWKEKRP 65
 Db 264 TPESLADLVSCWPEPSAPQELIRVPERLFTIDGDELKPSFHDQGGPWCCLWKEKRP 323
 QY 56 TEESMQLLPSWPEPSAPQELIRVPERLFTIDGDELKPSFHDQGGPWCCLWKEKRP 125
 Db 324 GGLLSRTPKRLKASASLITTPVLEKLOHPRHVEILGSEAEKRYFYKPHN 383
 QY 126 NAFQAGVYVNDNEPLFTMCWPLVWVATCLQQGLKSLLELQTSKTTAVVNL 185
 Db 384 NEQAGAFNLIGENEVLTMCFLPLVWVLTGLKQMETGKSLAQTSKTTAVVNL 443
 QY 186 LSLMQKPGAPRLQPPNGRGCSLAQSLWNQKILFEQDLKRGGLGEGDVSALN 245
 Db 444 SLLQSGSIEELHFLPSYLGSLAAGIWNQKILFEQDLKRGGLGEGDVSALN 503
 QY 246 FQKDKNSERF 256
 Db 504 FQKDKNSERF 514
 RESULT 10
 AAQ172670
 ID AAQ172670 standard; Protein; 1397 AA.
 XX
 AC AAQ172670;
 XX
 XX 31-MAY-2001 (first entry)
 XX Human NB-ARC and CARD containing protein (NAC) gamma isoform.
 XX
 XX Human; NB domain; nucleotide binding domain; NB-ARC and CARD; NAC;
 XX caspase-associated recruitment domain; CARD; TIM-Barrel-like domain;
 XX cysteine aspartyl protease; apoptosis; cytokine production;
 XX cytokine receptor signaling; therapy; inflammatory disorder; sepsis;
 XX fibrosis; arthritis; cancer; adenocarcinoma; leukemia.

CC processes, are principal effectors of apoptosis. CARD-5 and CARD-7 are used for screening modulators that modulates apoptosis, cytokine production, CC cytokine receptor signaling and other cellular processes. They can act as an immunogen for the production of polyclonal and monoclonal CC antibodies. They can also be used to diagnose and treat inflammatory disorders such as sepsis, fibrosis and arthritis and cancer pathologies such as adenocarcinomas and leukemias.

XX Sequence 1424 AA;

Query Match 31.9%; Score 759; DB 22; Length 1424;
Best Local Similarity 37.6%; Pred. No 236-68;
Matches 193; Conservative 71; Mismatches 165; Indels 69; Gaps 9;

QY 9 SMGLIFSWPESAPLOQLHVRERLLFIDGFEELKPSHPDGPVCLWEKRPTEL 68
DB 379 SLAEELKQSTATPAPRIQLSRERLLFIDGFEELKPSHPDGPVCLWEKRPTEL 68
QY 69 LLNSLIRKLLPELSLILITRTALEKLRHLEHRRHVEILGFSAPERKEYFYKHNAE 128
DB 439 LLSSLSLCKLILPEASEFLITARTALQNLPSUEQARWVEVLGFSSESRKEYFYRTDER 498
QY 129 CAGQVNVYRNEPFTMCVPLWCVVTCVCCGQARHLLPQTSRTTAVVYLYLSL 188
DB 499 QAIRAFVYKSKELWALCVVWVWKNIACTVYQVYKKEKLTITTCOLYLAQA 559
QY 189 MQPKPGARLPFPYVQRIQSLDADGLWQKILFEEDLERNKGLGSDVSAFLNNVFO 247
DB 559 LCAQPLGROL -----RLGSLAAEYVCKKTLFPIELRHSLDGLITFLNGILOE 612
QY 248 -----KQIND-----ERS-----PLATSP 262
DB 613 HPPLVSFPIELCPERFAANSYVLEDEKVRKSNVCTILEKTEAVNGHGLEASTR 672
QY 263 EEFMLAKETSHLEKSKWVSHPRVCLLWQCKQKAGSNHQQSLRPPFLYELIC 322
DB 693 ELGLGDSRGREVENIFRPIQVQVGNLWVWISQVLIIPHSBSNCTIVETR 726
QY 323 EEFELQALSHPOVAVENIAKEMHYVSNEDKQESACVIEHVNATYSADGERARCS 382
DB 727 AKTELQVMAHEEFGMCVETQNSHLLCTRCKPSRIVAKLOLI-----EGRQHS 779
QY 383 AGAHTLVQLPERTVLDAYSHHAAALTNFNLIEHEDFNALISRGVKKLCCGLRHEN 442
DB 798 WASTPMV LFWWEVTDAYWKLPSVIVKTVENIKELFLANSTSHRAVKSCHTIERR 814
QY 443 CKLQNL 449
DB 817 CULIELR 843

RESULT 12
AAB62571
ID AAB62571 standard; Protein; 1429 AA.
XX AAB62571;
XX

XX 23 JUL 2001 (first entry)
XX Human CARD-7 polypeptide.
XX

XX CARD-7; CARD-8; CARD-5; caspase recruitment domain; tumor; human;
KW autoimmune disorder; antiinflammatory; immunosuppressive; antia-leish;
KW antibacterial; antiviral; gene therapy.
XX

XX Homo sapiens.

XX Key Location/Qualifiers
FH Domain 329..645
FT /note= "nucleotide-binding domain"
FT Domain 333..341
FT /note= "kinase 1A subdomain"
FT Domain 424..413

FT /note= "kinase 2 subdomain"
FT 454..463
FT /note= "kinase 1A subdomain"
FT 475..482
FT /note= "motif 2 domain"
FT 487..494
FT /note= "leucine-rich domain"
FT 496..483
FT /note= "leucine-rich domain"
FT 484..481
FT /note= "leucine-rich domain"
FT 484..481
FT /note= "leucine-rich domain"
FT 483..493
FT /note= "leucine-rich domain"
FT 421..448
FT /note= "leucine-rich domain"
FT 458..476
FT /note= "leucine-rich domain"
XX MO200130813 AL
PN 03-MAY 2001.

PP 27 OCT-2002; 250260-US29746.

XX 27 OCT-1999; 90US 0428252.

XX (WILL) : MCGLENNIX PHARM INC.

XX Bertin C;

XX WFI; 2001-34740/36.

DR N-PSOE; AAF93651.

XX Identifying a modulator of interaction between caspase recruitment domain (CARD) 7 and CARD-5, for treating autoimmune disorders, the compound comprising the binding of CARD 7 and CARD-5 in the presence of the compound

XX Disclosure: Fig 1A-F; 83pp; Eng.165.

XX The invention relates to identifying compounds, that modulate interaction between caspase recruitment domain (CARD) 7 and CARD-5. The method involves measuring the binding of CARD 7 and CARD-5 in the presence of the compound (an increase in the binding of CARD-7 to CARD-5 in the presence of the compound compared to the binding in the absence of the compound indicates that the compound is a modulator of CARD-7-CARD-5 interaction). Modulators of CARD 7 and CARD-5 expression or activity can be used to treat or diagnose disorders such as cancers, bacterial or viral infections, autoimmune disorders (systemic lupus erythematosus, rheumatoid arthritis or arthritis), inflammatory disorders, organ-specific autoimmune including multiple sclerosis, Hashimoto's thyroiditis, or Grave's disease, psoriasis, graft rejection, allergies. CARD-7 and CARD-8 are useful as modulating agents in regulating a variety of cellular processes including cell growth and cell death. The present sequence represents the human CARD 7, an intracellular protein.

XX Sequence 1429 AA;

Query Match 31.9%; Score 759; DB 22; Length 1429;
Best Local Similarity 37.6%; Pred. No 236-68;
Matches 193; Conservative 71; Mismatches 165; Indels 69; Gaps 9;

QY 9 SMGLIFSWPESAPLOQLHVRERLLFIDGFEELKPSHPDGPVCLWEKRPTEL 68

DB 379 SLAEELKQSTATPAPRIQLSRERLLFIDGFEELKPSHPDGPVCLWEKRPTEL 68

QY 69 LLNSLIRKLLPELSLILITRTALEKLRHLEHRRHVEILGFSAPERKEYFYKHNAE 128

DB 439 LLSSLSLCKLILPEASEFLITARTALQNLPSUEQARWVEVLGFSSESRKEYFYRTDER 498

QY 129 CAGQVNVYRNEPFTMCVPLWCVVTCVCCGQARHLLPQTSRTTAVVYLYLSL 186

DB 499 QAIRAFVYKSKELWALCVVWVWKNIACTVYQVYKKEKLTITTCOLYLAQA 559

CC treating, or ameliorating a neural disorder or a disorder related to
 CC aberrant apoptosis modulation (either directly or indirectly), renal
 CC disorder, immune disorder (e.g. arthritis, asthma, acquired
 CC immunodeficiency syndrome (AIDS), or rheumatoid arthritis),
 CC haematopoietic, metabolic, reproductive, pulmonary or cardiovascular
 CC diseases, hyperproliferative disorders, neurodegenerative diseases
 CC (e.g. Alzheimer's disease, Parkinson's disease or Huntington's
 CC disease), developmental disorders, non-infectious disorders, nervous
 CC system diseases and/or disorders, and autoimmune diseases (e.g.
 CC Addison's disease, capillary anaemia, antiphospholipid syndrome, or
 CC allergic encephalomyelitis); the polypeptides are also useful as
 CC chromosome markers, for chromosome identification, gene therapy, and in
 CC identifying organisms from minute biological samples. This is the amino
 CC acid sequence of a leucine rich repeat containing protein.
 XX
 SQ Sequence 1429 AA;

Query Match 11.9%; Score 159; DB 23; Length 1429;
 Best Local Similarity 37.6%; Pred. No. 2, 28-68;
 Matches 183; Conservative 71; Mismatches 165; Indels 68; Gaps 9;

QY 189 MQPKFGAPLQPPNCRGLCSLAAGLAKXKLFEEQILRKFLEGGDVSAFLNMNIFQ- 247
 DB 559 LQAPLQGLQ- 612
 QY 248 KOINC- EPS- FLATSR 262
 DB 613 HPIPLSYSPHLCPOEFTAAASYVLEDEKQKGNKNCILIEKTLKAYGHNGLFSASTTR 672
 QY 263 FLUGLNEETRSHLEKSLCKWYSPIHKVTLCKTQSKAGSDSTLQOGLSLEFFSCLYEIQ 322
 DB 673 FLUGLSDEGEEMENIPIKRLSQ--GRNLMGWVPSQ--LLIQPHSLES.HCLVETR 726
 QY 323 REEPLCCALSHFQVIVWNSIASKWEHVSSEFLCKKCRSAQVCHLYGATYSADGSDRACS 382
 DB 727 NKTETLQWMAHEENGVC VETDMELLVCTCFIKFSHVKKLQL- 777
 QY 383 AGAHTLLVQLPRTVLLDAYSEHLAAALCTNPNLELSLYRNALGSRGVKLLCOGLRHN 442
 DB 778 TWSPTKWV LFRWVHVTDAYWQILPSVIVKVTNKLKESLSNLSLSHSAVKSUCKTLRPR 836
 QY 443 CKLQMLR 449
 DB 837 CLETLR 843

RESULT 14
 ABG78472
 ID ABG78472 standard; Protein; 1429 AA.
 XX
 AC ABG78472;
 XX
 DT 15-NOV-2002 (first entry)
 XX
 DE Leucine-rich repeat (LRR) domain containing protein #2.
 XX
 KW Human; human leucine rich repeat small intestine 1; HLRRS1; asthma;
 KW proliferative disorder; gastrointestinal disorder; renal disorder;
 KW neural disorder; reproductive disorder; calcium regulation; apoptosis;
 KW immune system; anaemia; human immune deficiency virus; HIV; cancer;
 KW blood coagulation disorder; autoimmune disorder; allergic reaction;
 CC

Example 1: Page 392-397; 415pp; English.
 PS
 CC The invention describes nucleic acids encoding human leucine-rich repeat
 CC containing proteins expressed in nervous system tissues, HLRRS1. The
 CC HLRRS1 polypeptide or the polynucleotide is useful for preventing,

CC treating, or ameliorating a neural disorder or a disorder related to
 CC aberrant apoptosis modulation (either directly or indirectly), renal
 CC disorder, immune disorder (e.g. arthritis, asthma, acquired
 CC immunodeficiency syndrome (AIDS), or rheumatoid arthritis),
 CC haematopoietic, metabolic, reproductive, pulmonary or cardiovascular
 CC diseases, hyperproliferative disorders, neurodegenerative diseases
 CC (e.g. Alzheimer's disease, Parkinson's disease or Huntington's
 CC disease), developmental disorders, non-infectious disorders, nervous
 CC system diseases and/or disorders, and autoimmune diseases (e.g.
 CC Addison's disease, capillary anaemia, antiphospholipid syndrome, or
 CC allergic encephalomyelitis); the polypeptides are also useful as
 CC chromosome markers, for chromosome identification, gene therapy, and in
 CC identifying organisms from minute biological samples. This is the amino
 CC acid sequence of a leucine rich repeat containing protein.
 XX
 SQ Sequence 1429 AA;

Query Match 11.9%; Score 159; DB 23; Length 1429;
 Best Local Similarity 37.6%; Pred. No. 2, 28-68;
 Matches 183; Conservative 71; Mismatches 165; Indels 68; Gaps 9;

QY 379 SLAEILGKDFATPAPIQQLSRREPILFLDGVDEPGWVLEQSSSELCLHWSQPADA 436
 DB 69 LNSLIRKLLPELSLITTRPTALFKLRRLEHPRHVEILGFSEAEKKEYFYKFNAAE 128
 DB 439 LLGSLGKTLPEASFLTANTTALQNLPSLEQARAVEVLGFSESSKKEYFYFYTDEP 498
 QY 129 QASQVNYVRNNEPLFTWCFVPLVGVWVVTCLQOLEGGLRGTSTTTAVVMYLLSL 188
 DB 499 QAIRAPFVYSKKEWALCVFVWVSWLACTCLNQOMKEKELTSLTTLCHYLAQA 558
 QY 189 MQPKFGAPLQPPNCRGLCSLAAGLAKXKLFEEQILRKFLEGGDVSAFLNMNIFQ- 247
 DB 559 LQAPLQGLQ- 612
 QY 248 KOINC- ERS- FLATSR 262
 DB 613 HPIPLSYSPHLCPOEFTAAASYVLEDEKQKGNKNCILIEKTLKAYGHNGLFSASTTR 672
 QY 263 FLUGLNEETRSHLEKSLCKWYSPIHKVTLCKTQSKAGSDSTLQOGLSLEFFSCLYEIQ 322
 DB 673 FLUGLSDEGEEMENIPIKRLSQ--GRNLMGWVPSQ--LLIQPHSLES.HCLVETR 726
 QY 323 REEPLCCALSHFQVIVWNSIASKWEHVSSEFLCKKCRSAQVCHLYGATYSADGSDRACS 382
 DB 727 NKTETLQWMAHEENGVC VETDMELLVCTCFIKFSHVKKLQL- 777
 QY 383 AGAHTLLVQLPRTVLLDAYSEHLAAALCTNPNLELSLYRNALGSRGVKLLCOGLRHN 442
 DB 778 TWSPTKWV LFRWVHVTDAYWQILPSVIVKVTNKLKESLSNLSLSHSAVKSUCKTLRPR 836
 QY 443 CKLQMLR 449
 DB 837 CLETLR 843

RESULT 14
 ABG78472
 ID ABG78472 standard; Protein; 1429 AA.
 XX
 AC ABG78472;
 XX
 DT 15-NOV-2002 (first entry)
 XX
 DE Leucine-rich repeat (LRR) domain containing protein #2.
 XX
 KW Human; human leucine rich repeat small intestine 1; HLRRS1; asthma;
 KW proliferative disorder; gastrointestinal disorder; renal disorder;
 KW neural disorder; reproductive disorder; calcium regulation; apoptosis;
 KW immune system; anaemia; human immune deficiency virus; HIV; cancer;
 KW blood coagulation disorder; autoimmune disorder; allergic reaction;
 CC

KW inflammatory condition; cardiovascular disorder; ischaemia;
KW neurological disorder; infectious disease; cytokine production;
XX expressed sequence tag; EST.
GS Unidentified.
XX WC26026:066-A2.
XX 06-AUG-2002.
XX 20-DEC 2002; 2001KW-US49739.
XX 22 DEC 2000; 2000US-257774P.
XX (BRIM : BRISTOL-MYERS SQUIBB CO.)
XX Feder J, Ramanathan C, Mintier G;
XX WPI: 2002 61952/66.
XX New isolated nucleic acid molecules encoding HLRBM1 polypeptides, or
PI their fragments and homologues, useful for preventing, treating and
PI ameliorating medical conditions, e.g. proliferative, gastrointestinal,
PI or renal disorders
XX
XX Disclosure: Page 229-233; 33pp; English.
XX The invention relates to isolated nucleic acid molecules (i) encoding
CC human leucine-rich repeat small intestine 1 (HLRBM1) polypeptides.
CC The nucleic acid molecules and polypeptides are useful for preventing,
CC treating and ameliorating medical conditions, such as proliferative,
CC gastrointestinal, renal, neural, or reproductive disorders; or disorders
CC related to aberrant calcium regulation or apoptosis modulation, either
CC directly or indirectly. They are also useful for treating, preventing
CC and/or diagnosing diseases, disorders and/or conditions of an immune system
CC by activating or inhibiting the proliferation, differentiation, or
CC mobilisation of immune cells; haematopoietic cells e.g. thrombocytopenia,
CC anaemia, immunologic deficiency syndromes, e.g. human immune deficiency
CC virus (HIV) infection, HTLV-BLV infection, blood coagulation disorders,
CC e.g. arterial thrombosis; autoimmune disorders, e.g. Addison's disease,
CC myasthenia gravis; asthma or allergic reactions; inflammatory
CC conditions, e.g. chronic prostatitis, sepsis, proliferative disorders,
CC e.g. cancer; cardiovascular disorders, e.g. arrhythmia, myocardial
CC ischaemias, aneurysms; neurological disorders, e.g. Alzheimer's disease,
CC Huntington's chorea; infectious diseases, e.g. measles, mumps,
CC pneumonia, or viral, bacterial, and fungal infections. The HLRBM1
CC polypeptides are useful for modulating cytokine production, antigen
CC presentation, or other processes such as host anti-immune responses.
CC ARG74454-ARG74474 represent HLRBM1 amino acid sequences and related
XX amino acid sequences of the invention
XX
SQ Sequence 1429 AA;
Query Match 31.9%; Score 769; DB 21; Length 1429;
Best Local Similarity 37.6%; Pred No 3,20,64;
Matches 183; Conservative 71; Mismatches 167; Indels 68; Gaps 9;
CY 9 SNQJIFSCWPPSPAPLDELIVYRRLFLITGDEDEKSPEDQGPCLWEKRRTEL 68
DB 179 SLAELIGKDGATPAPIRQIRSRPRLFLITGVDEPFWLQFTSSLECLFWSQPADA 438
CY 69 TANSUIRKKLAPLSLITTRTAEKIRHGLHPRFVEILDFSAERKEYFYKFNAAE 126
DB 439 LIGSLGKLTLEASFLITARTALQNI:PSLSEARWVEVLGFSSESKRYFYRFDER 498
CY 129 QAGQVNVYRNEPFTMCVFLPVQWVYTC:CCQEGGELLKQTSPTTAVVWLY:LSL 188
DB 499 QARAPFLVKSNEKLWALCVVWVSKLACTC:WQMKRKEK:TLTSKTTTLCFLYLAQA 558
CY 189 MGKPKGAPLPQPNQKGGCSAAAGLWKNQKILFTEQDLARKHGLGQDVSAFLNNIFQ 247
DB 559 LQAQPLGPGU:-----RDLSLAAGGQWQKTKLPSDPLRQKGLGALIS:FLKXGILQE 612

CY 248 -----KD'NC-----ERS-----FLAITSR 262
DB 613 HP:PLSYSEIN:CFQSEFFAAMSVVLEDEKGRGKSNCIIDLEKTLBAYGIRGLFASITR 672
CY 263 FLGGJLNEHTPSRLKSLKWKVSPHFMELLOWISKAAOSDGT:QQGSLRFFSFFSYLIQ 322
DB 673 FLGLGLSDGGERMEN:FCIRLSQ GRNIMQWVPS:Q-----LLQPHSDESILHCLVETR 726
CY 323 EEPFIOCALSHRFQV:VWNIAKSMENHYSFCLKRCBSACVLUHYGATYSADGEARCS 382
DB 727 NKTETIQNAHFEKEMC:VETQNHLCVCTPCKFSRHVKKQLI-----EGRQHS 777
CY 383 ASANTLVCQPEPTVILDAYSEHIAAANTNRLIELSLYRNALSGRGVKLQCGGLRHPN 442
DB 726 TNSPTWV LKRWVPVDAWQ:LFSLKVTENLKEILDGNSLSHSAVKS:CKLRLRPP 816
CY 443 CKLQNR 449
DB 817 CLETER 94;
RESULT 15
ABR77916
ID ABR77916 standard; protein; 1429 AA.
XX ABR77916;
XX AC ABR77916;
XX 07 OCT-2002 (first entry)
XX Human leucine rich repeat domain containing protein KIAA0926.
DE Human leucine rich repeat domain containing protein KIAA0926.
XX Human leucine-rich repeat; HLRBM1, proliferative disorder;
KW immune condition; apoptosis; signal transduction; autoimmune disease;
KW haematopoietic cell disease; graft versus host disease; allergy; asthma;
KW cardiovascular disorder; neurological disease; pheromone;
KW pulmonary disease; chronic obstructive pulmonary disease;
KW allergic rhinitis; bronchial hyperresponsiveness; reproductive disease;
KW haematopoietic disease; platelet disorder; Bernard-Soulier syndrome;
KW inflammatory disorder; systemic lupus erythematosus;
KW cardiovascular disease; Cancer; KIAA0926.
XX Homo sapiens.
XX WC200252011-A2.
XX 04 JUL-2002.
XX 20-DEC 2001; 2001KW-US49740.
XX 22-DEC-2000; 2000US-25774P.
XX (BRIM : BRISTOL-MYERS SQUIBB CO.)
XX Feder J, Ramanathan C, Mintier G;
XX WPI: 2002-566676/60.
XX New HLRBM1 nucleic acids for preventing, treating or ameliorating e.g.
PI proliferative disorders, immune conditions, a disorder related to
PI aberrant apoptosis modulation or developmental disorders.
XX
XX Example 1: Page 362-366; 37pp; English.
XX
CC The present sequence represents a human leucine-rich repeat domain
CC containing protein, which was used as a probe to search for leucine-rich
CC repeat containing protein HLRBM1. HLRBM1 polypeptides and
CC polynucleotides are useful for preventing, treating or ameliorating a
CC medical condition such as a proliferative disorder, immune condition, or
CC a disorder related to aberrant apoptosis modulation, either directly or
CC indirectly, and in modulating signal transduction activity in various
CC cells, tissue and organisms. They are also useful for treating,
CC preventing, or diagnosing diseases of haematopoietic cells, autoimmune
CC disease, graft versus host disease, allergic conditions (e.g. asthma),

cardiovascular disorders, and neurological diseases, and for increasing the organ's ability to synthesize and/or release pheromones. The polypeptide may also be used in treating, preventing or ameliorating pulmonary disease (e.g. chronic obstructive pulmonary disease, allergic rhinitis, or bronchial hyperresponsiveness), reproductive disease, haematopoietic disease, platelet disorders (e.g. Bernard-Soulier syndrome), non-infectious disorders (e.g. innate immunity to bacterial pathogens, or adaptive immune response), immune and inflammatory disorders (e.g. systemic lupus erythematosus), cardiovascular diseases and cancers. HRRRM1 nucleic acids may further be used in chromosome identification or mapping, as a chromosome marker, as molecular weight markers, as diagnostic probes, in gene therapy, in raising anti-DNA antibodies, or as antigens for eliciting immune responses.

XX	Sequence	1429 AA;
QY	Query Match	31.9%; Score 759; DB 23; Length 1429.
DB	Best Local Similarity	37.6%; Pred No. 2,20,68;
	Matches 103; Conservative 7; Mismatches 169; Indels 69; Gaps 9;	
QY	9	SMQDLIFSCWPEPSAPLOQSLIRVPERLLFIDGPRDLKPSRDPGGWGLCWEEKRPTL 68
DB	179	SLAELICKGGTATPAPIRQILSRPERLLFIDGVPEQWVLQPSSEGLHWSPQPCA 438
QY	69	LNSLIKKLLDELILLITPTALEKURHLERPRHVELGFSAEKPKSYKYVHNAE 128
DB	439	LSGLIGKTLIFEAGFLITARTALQNLSPSLQARWVEVGFSSSRKEYFYFDER 496
QY	129	QARQVENVYDRKEPLFTMGFVPLVWVYCTIQQLQGLSGHLIRQTSRTTAVYMLYLLS 188
DB	499	QAIRAFRLVSKKELWALCLVWVSWLACTLQXQKKKKELTLISKTTTTLCHYLAQA 558
QY	189	MQPKGAPESQPPKQRGGLCSAAFGLNQKILFEQQLRKIGLDGDSVAFKNNIFQ- 247
DB	559	LOAQRLGPQ---RDICSLARGIWKQKTLFSPDLKXHGIGCAISTFLKMGIIQE 612
QY	248	-----KQING-----ERS-----FLALISR 262
DB	613	HPDPLSYSTIHCFUEFFPAKSYVLEDEKGRKSNCTHLEKTLKAYGIHGLFGASTTP 672
QY	263	FLFGLNEETRSRLKSLCWKVSPIKPILOWIKSKAQSDGSTQQQSEDEFFSCLYEIQ 322
DB	673	FLGLSLDGEKEXENIFHCRLSQ-GRNLMQWVPSLQ-LLQPSIESLHCLYETR 726
QY	323	SEEFQCALSHFQVIVVSNIAKNNKRVVSGPDKRTPRAVYVHYGATYASDGEGRAPCS 382
DB	727	KATILTQVAHFEKNGC VETDMELVLTETIKPSRAVYVGLI---EGGQHS 777
QY	383	AGAHLLVLQPERTVLDAYSSEHMAAATNNILIELNLYPKALDSRGVKLCCGLRHPN 442
DB	778	TWSPNVV-LFRWVPTDAYWQILFSLVAVTFLRELKSNSSCHSAVKSLSKXTIRPR 836
QY	443	CKLQNR 449
DB	837	CLETLR 843

Search completed: October 31, 2003, 07:04:04
 Cpu time : 85 secs

GenCore version 5.1.6
Copyright (c) 1993-2003 Computer Ltd.

QM protein - protein search, using sw model

Run on: October 31, 2003, 03:53:49 : Search time 44 seconds
(without alignments)
338,632 Million cell updates/sec

Title: US-10-028-374-2
Perfect score: 2377
Sequence: 1 MQSATEKSYQDLFSWFE.....HWKMLGQGRHFNKQLNR 449

Scoring table: RLOSM62
Gapop 10.0, Gapext 0.5

Searched: 283309 seqs, 96168662 residues
Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum March 120%
Listing first 45 summaries

Database: PIR 75%
1: PIR1
2: PIR2
3: PIR3
4: PIR4

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	759	31.9	1192	2 T12255	hypothetical prote
2	635	22.5	1171	2 A56118	water protein (imp
3	327.5	13.8	483	2 S27862	Maspressin recepr
4	214	9.0	1130	2 A48473	MHC class II trans
5	150.5	6.3	1004	2 T12255	hypothetical prote
6	118.5	4.9	1078	2 T12255	hypothetical prote
7	113	4.8	456	2 A11457	ribonuclease inhib
8	110	4.6	456	2 S27862	ribonuclease inhib
9	109	4.6	388	2 S27862	cyclin B - startis
10	107	4.5	461	2 A31958	ribonuclease-angi
11	106.5	4.5	1232	2 A56478	neuronal apoptosis
12	106	4.5	1010	2 T36363	probable large ATP
13	102	4.3	2110	2 F96503	unknown protein TS
14	99	4.2	506	2 A49541	r complex-associat
15	96.5	4.1	1172	2 S51623	cut14 protein - fi
16	95	4.0	2021	2 A23267	serine/threonine k
17	93	4.0	2747	2 B49112	fat facets (faf) s
18	94.5	4.0	1447	2 T42628	neuronal apoptosis
19	94	4.0	518	1 A47626	glycine hydroxymet
20	93.5	3.9	853	2 S58175	ary. hydrocarbon r
21	93	3.9	1277	2 B84517	hypothetical prote
22	91.5	3.8	459	2 A46312	transmembrane PKB5
23	91.5	3.8	613	2 S74611	hypothetical prote
24	91.5	3.8	613	2 F89301	conserved hypothet
25	91	3.8	1306	2 T55178	hypothetical prote
26	90.5	3.8	797	2 D86269	hypothetical prote
27	90.5	3.8	1402	2 T17456	cell surface prote
28	90	3.8	394	2 T19218	hypothetical prote
29	89.5	3.8	2529	2 T20119	hypothetical prote

ALIGNMENTS

RESULT 1

T12255
Hypothetical protein DKFZps86G1822.1 - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 15-Oct-1999 #sequence revision 15-Oct-1999 #text_change 15 Oct-1999
C:Accession: T12255
R:Koehrer, K., Beyer, A., Vewes, H.W., Gassenhuber, G., Wiemann, S.
submitted to the Protein Sequence Database, September 1999
A:Reference number: Z19122
A:Accession: T12255
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1192 x 45
A:Cross-references: EMBL:AL117470
A:Experimental source: adult uterus; clone DKFZps86G1822
C:Genetics:
A:Note: DKFZps86G1822.1

Query Match: 31.9%; Score 759; DB 2; Length 1192;
Best Local Similarity 37.6%; Pred. No. 43e-57;
Matches 183; Conservative 71; Mismatches 165; Indels 66; Gaps 9;

Cy	9	SMQDLFCWFRNSAFLOQ:RYPERLLFIIDFELKPSFDQPOQPCWLOWEKRTPEL 68
Db	98	SLAEILFTHTATPAITROTLSKPERLLFIIDGVDPQWVQEPSELCLKRSQCPADA 157
Cy	69	LKASLIERKMLRPSLSLITRTPLAEKILPILFHPRHVEILGFSEAEKKEYFYKHNAE 128
Db	156	LAGSLIKNTIDPASFTTARTALCNLPSIQARWVEVLGFSSSRKEYFYPTDTER 217
Cy	129	QAGQVNYVNEPELFTYGFVPLVWVWVTCQQQLEGGLRQTSRTTAYVMYKLSL 188
Db	218	CAIRAFELKSKNELNALCNVWVWSMLATCQMKRKEKTLTKITTTTCLHYLQA 277
Cy	189	MQPQKAPRIQPPFNKSGLSLAACGLWVKKILFPEQDLRKGLGDCVSAFLWNKIQ- 247
Db	278	LAQCP:SPQL --- PCLSLAENGIWKKTKTSPFDCLRKHGLGCA:ISTFLKYGILQE 331
Cy	248	-----KQINC-----ERS-----EALATSR 262
Db	332	HPILKSYNFIHLGQFQFFFAKSYVLEDEKGRKUNCIIDLEKTLKLEAYGIHGLFASITR 391
Cy	263	FLGLNEETFRSLEKSLCWKYSPIHKMLDLWIOSKAGSDGSLTQQSLEFFSCYEQ 322
Db	392	ELGLLRLDEGEREMENIFNCLISQ--GRNCKWQVPSLQ----LLLPMSLSLHCLCYETR 445
Cy	323	EEFFICGLSHFPVIVVNSIAKKNHVSFCLKCRSAQVHLVGYATYSADCEPARCS 382
Db	446	KKFLQVWVHFFENKVC VETVELVATFTFRFSRVKKLQLI-----EQRQRS 496
Cy	383	AGAPTLLVQVHSTVGLDVAISHLAAGCTNPKNFELSYRNALSGVKLCGGLRHPK 442

vetispiradiene syn
hypothetical prote
probable exodeoxyr
giant protein p619
ABC transporter ih
malt regulatory pr
heterocyst envelo
hypothetical prote
protein F28Q1.3 (1
hypothetical prote
hypothetical prote
ATP-dependent dscN
hypothetical prote
probable chromatin
hypothetical prote

DQ 497 TWSPTVVV-LPRKVVTDAYWQILSFVIVKTRNRELISKSLSHSAVKSUCKTILRRP 555
 QY 443 CKIONLR 449
 DQ 556 CUEETLR 562
 RESULT 2
 A59000
 matrix protein (imported) - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 03-Aug-2001 #sequence, revision 03 Aug 2001 #text change 03-Aug-2001
 C:Accession: A59000
 K:Long, Z.B.; Nelson, J.M.
 Endocrinology 140, 3726-3726, 1999
 A:Title: A mouse gene encoding an oocyte antigen associated with autoimmune premature ov
 A:Reference number: A59000; PMID:99360614; PMID:10433232
 A:Accession: A59000
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1111 <KUR>
 A:Cross-references: GB:AF074018; NID:9402692; PID:AA017421; PID:95802698
 C:Genetics:
 A:Gene: Muter
 Query Match 22.5%; Score 51.7; DB 2; Length 1111;
 Best Local Similarity 29.1%; Pred. No. 1.2e-37;
 Matches 143; Conservative 85; Mismatches 209; Indels 56; Gaps 6;
 QY 9 SYGCLIFSCWKPSPAPQGLIVRPFSLAPLAPFLKSLRFDPSGPMILWEEKRPTEL 68
 DQ 241 SLACIAKPCPSWNVTKMSJPERALVIAIKMSVLCHEMTLSRWKKEQPIYI 300
 QY 69 LANSILKKLPELSLITTEETALKEHLHFNHKEILNPSHAKKEVEKYFNAAE 128
 DQ 301 LMYSLKALKLPQSLITTRATGLEKLSMVSGLYILVESLSASRSQLENISHES 360
 QY 129 CACQENYVRDEPLTMEVEVLVWVYTCQCCQCEPFLIPUSRTTAVY 185
 DQ 361 DRIOVSHSIEHQLPQQCAPSVSLVGEALGLQRKIGAKHLRQCGLISLYATVPHQ 420
 QY 186 LSLMQPKGAPRIQPPNKGKLSAANDGLNKKLIFRQELERHGLSGEDSAPENMKI 245
 DQ 421 LILKRSQSALSQSEQITLVGLVYVALEKRWINGVAVVINGLKNYSKSESELALFHNKI 480
 QY 246 FQKDNERS 259
 DQ 481 LLOVHNSQVYVSHSLQCTEAAQVYVTELELWNSJKECTEAGKIMEVKTQTRL 540
 QY 260 TSREPLGLAREETHSHLEKSLAKVNSDKNVYGLQWQLKACQUGSTIQGSLFEFSC 317
 DQ 541 LKXKRLFGSKMKILKTLTEVFEVETVIVIVKQJHVSLLAQVNGTSPKTDLEAFYC 600
 QY 318 LYETQETEQALSFPQVTVNENASRMEHVSEFTLKEKQAGVILHVGATYSGDGD 379
 DQ 601 LRFSCQSEFVGKALRFQEWLLINQAVMLKVSAYGLQKQKRAIVLRQLLSVNT 659
 QY 378 PARCSAGMTLLVQLPERTVLLANSEHAAALCNPHLELSLERNMAGRGVLLQCG 437
 DQ 660 LELCP 715
 QY 438 LRHPNCKLQNL 448
 DQ 716 LRNQSRIOQL 726
 RESULT 3
 S2788C
 Vasopressin receptor rat
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 10-Apr-1993 #sequence, revision 10-Apr-1993 #text change 05-Nov-1999
 C:Accession: S2788C
 R:Herrera, V.L.; Ruiz-Opazo, N.

submitted to the EMBL Data Library, February 1992
 A:Description: Characterization of a cDNA encoding an AII and AVP receptor isoform.
 A:Reference number: S27880
 A:Accession: S27880
 A:Molecule type: mRNA
 A:Residues: 1-483 <HER>
 A:Cross-references: ENEL:82518; NID:9202825; PID:AA036231; PID:9202826
 C:Genetics:
 A:Gene: AVP
 Query Match 11.8%; Score 327.5; DB 2; Length 483;
 Best Local Similarity 27.8%; Pred. No. 3.2e-20;
 Matches 109; Conservative 53; Mismatches 123; Indels 105; Gaps 11;
 QY 164 LGGGLRQTSRTTAVVWLYLLSNPKPGAPRLQPPNQGRLCSLAADGLMKNKILF 222
 DQ 1 MELGRELSATSKTTTSVYLFF.TSNLKSAGTNSPRVQ--GELRMLCLRLAREGILKHQAQF 58
 QY 223 EEOQLR-KHGLDGEDVSA-----PLNNI----- 245
 DQ 59 SSKDLERLKLKAGSQVQVYFLSKKELPGVLETVVTVYFIDQSFOEPLAALSYLDAEGA 118
 QY 246 -----PCKDINKT---RSFLATSRPFGLNSETRSHLEKSLCWKVSPLKVKYLLQW 295
 DQ 119 PENSASISVOMLNSDAHLRCHLALTIPTFGLLSTERIDIGNHFGCVVPEVKZCTLEK 178
 QY 296 IQSKAQSGSTG----- 330
 DQ 179 VQGSQKQVATVGAERKDELKDFEAEEREEEREEELNFGLELLYCLYETQEDDFVQA 238
 QY 331 LSHFGVTVSN: ASKVEHNVSSFCARCSQVHLVYGATYSADGEDRAE----- 380
 DQ 239 LSSLFEMVLRVRLTMDLEVLVYCVCCQPCQCALRVSGLVAAKEXKXKXKXKXKXKXKXKX 299
 QY 481 CSAGAL: TLVQCCPRTV-----LDDAYSEHLAAALTNPKLIELS 420
 DQ 299 GSSTCKQPHASLRPLSRAMITQQGSLITLSHCKIPDAVCRLSELKALVAPSLPELG 358
 QY 421 LYNALSRVRLICQGLRHPNCKLQNL 449
 DQ 359 LQGNRTEAGRLISQSLAMPKQKQVOTR 367
 RESULT 4
 A48843
 MHC class II transactivator human
 N:Alternate names: AIIA
 C:Species: Homo sapiens (man)
 C:Date: 06-Jan-1995 #sequence, revision 06 Jan-1995 #text change 08-Oct-1999
 C:Accession: A48843
 R:Steimle, V.; Otten, J.A.; Zufferey, M.; Mach, B.
 Cell 75, 135-146, 1993
 A:Title: Complement activation cloning of an MHC class II transactivator mutated in heredit
 A:Reference number: A48843; PMID:94026536; PMID:9402893
 A:Accession: A48843
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-1130 <SIE>
 A:Cross-references: GR:A74101; NID:9414112; PID:CAAS23541; PID:9414113
 Query Match 9.0%; Score 214; DB 2; Length 1130;
 Best Local Similarity 24.3%; Pred. No. 6.6e-10;
 Matches 118; Conservative 61; Mismatches 187; Indels 120; Gaps 23;
 QY 1 MNSATECSMQLIFSCWPEPSAPQCE---LIRVPERLJFIIDGFDELKPS---FHUPQ 53
 DQ 456 LKPPGAYGILQLFSLGQPLVAADAEVSHILKRFEDRVLLILDFAFELEAQQGFLHSC 515
 QY 54 GPATLAKWEKSEFTEL--LINSJIRKLLPELSLCTTRP-----TALKKLELLEHPRV 106
 DQ 516 GP--APAEFSLKLLAGLQKALRGCTLTATPRGRVLSLSKADAEF----- 564
 QY 107 ETLGSEAEKKEVEYFYFENA---EQAGGVPRVVDNEPLFTMCVPLVCLVYVTCQJQ 163

```

Db      565 ELGSPMECAQAYVYRYESSOMTEHQDAUHLKRNIFLLSHSHSPTLCRAVCQ--SEAL 624
QY      164 LEGGGLLQTSRTTAVVYLLYLLYKQKAPAPLQPPHQKQ--LCSLAAD GLKNCKIL 221
Db      645 HELGECALPS--TLTSHYVGLD-- --GFAALGSPHALASAKLAKELGRRHGSTC 675
QY      224 FEQ-- --DURKIGS-- --QJEDVASF-- --LKNVIFOKDINC 252
Db      616 QEDGFSADVRTWAKGLVYUHFRAESLAFSLFLLFELAKLALASLKKDEL-- 733
QY      253 ENSFLALT----- --SPLFLLNEF--ASHLKKLWKKVSH--KMLLOWIQS 298
Db      734 PQYUALTFRKRRYPDNWLEWVPEFADALFQFARVLSALL--GPSAAASVDRKKQVLA 790
QY      299 --KAQSDGST:QGSLSGFEFFSLVLCQEEFTQCALSHHFQVYVNSNTASKMEHWSSFC 355
Db      791 RYKRLQPGQTJPARQLLELHCAHAEAGTQCH-- --VWOELPGGJ-- --SFLG 837
QY      356 KRCSAQVHLHYGATYSADGGD----- PARCSAGAHITL-- --VQLPVRTVL 399
Db      838 TRLTPEDA--HVLGKALEMAGGCFSLQJSTJLCTGSLGSLVGLSCVTPFRAL--SDTVALW 896
QY      400 DAYSEH 405
Db      897 ESLPQH 522

RESULT 5
T31665
Hypothetical protein OS1.4 - sea squirt (Cliona intestinalis)
C:Species: Cliona intestinalis
C:Date: 29 Oct 1999 #sequence_revision 29 Oct 1999 #text_change 20 Jun 2000
C:Accession: T31665
R:Bird, A.P.; Clark, V.J.; Jones, S.J.; Lettgeb, S.; Tennard, N.; Tweedie, S.
submitted to the EMBL Data Library, October 1996
A:Reference number: Z11050
A:Accession: T31665
A:Status: preliminary; translated from GB/EMBL/DBCB
A:Molecule type: DNA
A:Residues: 1:1004 <B.R>
A:Cross-references: EMBL:Z82904; PDB:1CAB02589.1
C:Genetics:
A:Introns: 30/1; 539/1; 593/1; 662/1; 727/1; 769/1; 817/2; 861/2; 960/1

Query Match 6.33; Score 16.5; DB 2; Length 1075;
Best Local Similarity 19.53; Pred. No. 6.16;
Matches 99; Conservative 79; Mismatches 146; Indels 145; Gaps 23;

QY      20 EESAPJQELIRVPERGLTIDFHLIKRPHLPQWTLK--RKRITELLINSLRKK 77
Db      132 ELSKALITFLHNNPVA--FFDGLIFASTNEFARIPHCKLDGKSKPVDIMKN--LFDNL 188
QY      78 LPLSLSLITTPPTALEKHLLEHFRHVEILGPSAEKREKYFYKFNHNAEQGVNVR 137
Db      189 LIPAKIVVTSTPDMPNLQCYPTS--FEVLSTLEAKNNL-- --GTCLGCKYPAI 242
QY      138 R-- --DNEP-LFTMCPVPLVGVNVTCTGCGGLEGGILLRQTSRTTAVVYLLSLMQ-- 190
Db      243 KKLIDQPNLAHLCYLP--NFILVFCLLN--EGSDI-- --KMTQVLFSTNTRFVESH 296
QY      191 -----PKPGAPRIOPPNCRGSLAAG--KNCKILFEEQLKRLDGE----- 235
Db      297 LKGEVPLDVGAEVKK-----LACLAYKGIQKKLVFEKTFDQVKLADENVTHFT 349
QY      236 --DVSALNNM:FQDKINCESEF----- --LAKLALALASLKKDEL-- 733
Db      350 YVDLSSGIRIKILEN-- --KSYF--HHIQWQFYAAVYLMQVSVYREFEQ--KTIFEDTQWS 406
QY      259 JTSRELFQJN----- -- --RPTSHLSKSLKLVKVSHPHCKMDLQ W: 296
Db      407 VVVKFMEGICNPAYKQIKLIPATNIRKYEEKKFEESYWNESISAKSETQIPFESWL 466

```

```

QY      297 QSKAGSDST:QGSLSGFEFFSLVLCQEEFTQCALSHHF----- --YEIQEEFFIQCALSHF----- 334
Db      467 HE-- --VNCDET-- --SKREKQCLVGLKLVKHKH--PEVKOLVYALKSKTKKHKLRSKWT 520
QY      335 -- --QVLT-- -- --VYNTASKMEHWSSFC--KRCSAQVHLHYGATYSAD 374
Db      521 TTVETLETGKRGHHTTTTTRFVNNKLEMDSDM-- --ELLQHL----- --D 563
QY      375 GEDPARCAGAHITLVCQPERTVALLAYSEHAAALCINPNLIELSLY 422
Db      564 AVEKLEF-- -- --GVTHLSYNSLSINAINORSNKIOLQJW 599

RESULT 6
T31668
Hypothetical protein OS1.5 - sea squirt (Cliona intestinalis)
C:Species: Cliona intestinalis
C:Date: 29 Oct 1999 #sequence_revision 29 Oct 1999 #text_change 20 Jun 2000
C:Accession: T31668
R:Bird, A.P.; Clark, V.J.; Jones, S.J.; Lettgeb, S.; Tennard, N.; Tweedie, S.
submitted to the EMBL Data Library, October 1996
A:Reference number: Z21052
A:Accession: T31668
A:Status: preliminary; translated from GB/EMBL/DBCB
A:Molecule type: DNA
A:Residues: 1:1075 <B.R>
A:Cross-references: EMBL:Z82904; PDB:1CAB02589.1
C:Genetics:
A:Introns: 48/1; 557/1; 611/3; 685/2; 727/2; 739/1; 779/3; 820/3; 853/1; 914/2; 993/1

Query Match 4.93; Score 116.5; DB 2; Length 1075;
Best Local Similarity 19.53; Pred. No. 6.16;
Matches 98; Conservative 76; Mismatches 203; Indels 125; Gaps 19;

QY      20 EESAPJQELIRVPERGLTIDFHLIKRPHLPQWTLK--RKRITELLINSLRKK 77
Db      132 ELSKALITFLHNNPVA--FFDGLIFASTNEFARIPHCKLDGKSKPVDIMKN--LFDNL 208
QY      80 PELSALITTPPTALEKHLLEHFRHVEILGPSAEKREKYFYKFNHNAEQGVNVR 139
Db      209 PRAKIVVTSTIHMVKLHPDYRPTS--FEVLSTLEAKNNL-- --GTCLGCKYPAI:KK 262
QY      139 --DNEP-LFTMCPVPLVGVNVTCTGCGGLEGGILLRQTSRTTAVVYLLSLMQ-- 190
Db      263 LILQQPNLAHLCYLP--NFILVFCLLN--EGSDI-- --KMTQVLFSTNTRFVESH 316
QY      191 -- --PKPGAPRIOPPNCRGSLAAG--KNCKILFEEQLKRLDGE----- 226
Db      317 QVPLDVGAEVKK-- --LAPLAYKGIQKKLVFEKTFDQVKLADENVTHFTYV 369
QY      227 -----LRKFLGCHVNSAFANVNIPOKDCINCESEFALTSR----- 262
Db      370 QKSSGIRIKILEN-- --KSYF--HHIQWQFYAAVYLMQVSVYREFEQ--KTIFEDTQWS 428
QY      263 LPLSLSLITTPPTALEKHLLEHFRHVEILGPSAEKREKYFYKFNHNAEQGVNVR 315
Db      429 PVEGICNPAPKQK-- --LVFPATN:KQYEKKELM-- --VPMMSLSWARGDLIRF 481
QY      316 SCLYIQEEFTQCALSHFQV-- --IWSNIAKSYEHMVSSFC--KRCSAQVHL-- --YCAT 370
Db      482 GWLREYNLCSSKKPEDY:PVGLKVDAPKHLSEVKDLV-- --YALKSFTKPKHLRLDSVET 539
QY      371 YSACHEDPARSAGAHIT----- -- --LLVQLPE----- --RTVLADAYSE 404
Db      540 TTEVLETIRGVHGTITTTTTRFVNNKLEMDSDM-- --ELLQHL----- --D 599
QY      405 HRAALCINPNLIELSLY 426
Db      600 RLSNALNCSNKIQLV-- --KSKL 621

RESULT 7
A31857

```

A:Cross-references: EMBL:X16829; NID:93703; PDB:CAA34624.1; PID:g39702
A>Note: the sequence from Fig. 6 is inconsistent with that from Fig. 5 in having 12-AAs
A:Note: 109-Val was also found
C:Superfamily: cyclin
C:Keywords: cell cycle control

Query Match 4 ft Score 159 DB 2 Length 388;
Best Local Similarity 23.3%; Pos No. 0.17;
Matches 64; Conservative 43; Mismatches 16; Indels 52; Gaps 12;

QY 174 RTTAVLVLYSLHMKQKPGAPLPDPNKGCSLAADGLWNCKILFFEQDPRKESD 243
DB : : : : :
DB 63 SKATSLSSWGLN-VPEYKAKPCSPRYMSENSALEA-FSONCLEGVEDLCNDPD 120
QY 234 GRDY SAILNNIFKCNCTSRSHALTSRFLHLNBETRSLEKSLGWKVSHEIKMD 292
DB : : : : :
DB 121 KFCIAKSEFVN-CIYYNRKLREFAKVRIDYMTIGETIRVRSIL-IDMLVQVHLRFPL 176
QY 293 LGWI - - - - - GKAKSGSTLQGSLEFFESCYLE-QESEPQQALSHP 334
DB : : : : :
DB 177 LGSTLFLLICLDYLYEPVSKKKCLGVGV-----SMLIAKYEMYPPEIGDF 227
QY 335 QIVYSVMIAKMEHVSYSC-LKR-----CRSAQVLHYGATYSGGEUNARCSA 389
DB : : : : :
DB 228 WITDNAYTKAQTRSMCHNLRLPDSFKPKLC--IHFLRRSKAGRVFGQKHV 280
QY 384 SAHTLAVPFTSVLLDAVSEHAANCTNNILE 418
DB : : : : :
DB 281 AYVDELTPFAVFVPYSEIFAANAALGCCSKILE 315

RESULT 10
A31958
ribonuclease-angiotensin inhibitor - human
N:Alternate names: ribonuclease inhibitor, p.argenta.
C:Species: Homo sapiens (man).
C>Date: 0-Jun-1990 #sequence revision 07 Jun 1990 #text_change 19-May-2000
C:Accession: A31958; S2012; S23933; S48636; T47198
R:Lee, F.S.; Fox, E.A.; Zhou, H.M.; Strydom, C.G.; Vallee, B.L.
Biochemistry 27, 8645-8653, 1988
A>Title: Primary structure of human placental ribonuclease inhibitor.
A:Reference number: A31858; MJID:89118269; PMID:3219362
A:Accession: A31858
A:Molecule type: mRNA
A:Residues: 1-461 <LEE>
A:Cross-references: GRFM22414; NID:g18626; PDB:AAA59130.1; PID:g397040
R:Schnieder, R.; Schneider, Schuzter, E.; Thurnher, M.; Auer, B.; Schweiger, Y.
EMBO J 7, 4251-4259, 1988
A>Title: The primary structure of human ribonuclease/angiogenin inhibitor (RA) discal
A:Reference number: S2012; MJID:8920799; PMID:3243277
A:Accession: S2012
A:Molecule type: mRNA
A:Residues: 1-422/SE/425-461 <SCH>
A:Cross-references: EMBL:X1973; NID:G35843; PDB:CAA32151.1; PID:g35844
A>Note: Part of this sequence, including the carboxyl end of the mature protein, was c
R:Crevel-Grieffy, L.; Costerill, S.; Schullier, Z.
Biochim Biophys Acta 1122, 107-112, 1992
A>Title: Characterisation of a tryptic peptide from human placental ribonuclease inhibi
A:Reference number: S23933; MJID:92338217; PMID:1631192
A:Accession: S23933
A:Molecule type: protein
A:Residues: 174-195 <CRE>
R:Nadano, D.; Yasuda, T.; Takeshita, H.; Uchida, K.; Kishi, K.
Arch Biochem Biophys 312, 421-428, 1994
A>Title: Purification and characterization of human brain ribonuclease inhibitor.
A:Reference number: S48636; MJID:94311593; PMID:8037455
A:Accession: S48636
A>Status: preliminary
A:Molecule type: protein
A:Residues: 2-14 <NAS>
R:Poustka, A.; Neumannhuber, P.; Vewes, H.W.; Weil, B.; Wiemann, S.
submitted to the Protein Sequence Database, March 2000
A:Reference number: T24194

Search completed: October 31, 2003, 07:02:39
Job time : 5: secs

GenCore version 5.1.16
Copyright (c) 1993 - 2001 GenPharm Ltd.

GM protein - protein search, using sw model

Run on: October 31, 2003, 04:20:04 : Search time 25 seconds
(without alignments)
844,600 Million cell updates/sec

Title: US-10-028-374-2
Perfect score: 2377
Sequence: 1 MNQSAIEGSMQDIFSCWPE.....GVKILNQSLRHNP:KIONR 449

Scoring table: BLOSUM62
Gap: 10.0, Gapext: 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127543

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 38
Maximum Match 100
Listing first 45 summaries

Database: SwissProt_41*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	DB ID	Description
1	2334	99.2	1062	1	PVAF_HUMAN	P59046 rat
2	1157.5	48.7	1034	1	C1S1_HUMAN	Q94P20 homo sapien
3	1145	48.2	1033	1	C1S2_MOUSE	Q94P20 mus musculus
4	759	31.9	1473	1	NALI_HUMAN	Q95600 homo sapien
5	674.5	28.4	892	1	PVAF_HUMAN	P59046 homo sapien
6	671.5	28.2	843	1	PVAF_MOUSE	Q94P20 mus musculus
7	655	27.6	894	1	PVAF_RAT	Q94P20 rat
8	652.5	27.5	994	1	NALI_HUMAN	Q95602 homo sapien
9	564	23.7	1052	1	NALI_HUMAN	Q94P20 homo sapien
10	563.5	23.7	992	1	PVAF_HUMAN	Q94P20 homo sapien
11	556.5	23.4	1200	1	NALI_HUMAN	P59347 homo sapien
12	535	22.5	1111	1	NALI_MOUSE	Q94P20 mus musculus
13	428.5	19.0	1033	1	PVAF_HUMAN	P59046 homo sapien
14	261.5	11.0	1040	1	CARF_HUMAN	Q94C29 homo sapien
15	261	11.0	953	1	CARF_HUMAN	Q94C29 homo sapien
16	253	10.6	1020	1	CARF_MOUSE	Q84320 mus musculus
17	233	9.8	983	1	CARF_MOUSE	Q84320 mus musculus
18	214	9.0	1230	1	C2TA_HUMAN	P59376 homo sapien
19	174	7.3	1155	1	C2TA_MOUSE	P59376 mus musculus
20	117.5	4.9	1403	1	B1RA_MOUSE	Q94W55 mus musculus
21	113	4.8	456	1	RINI_PIC	P18775 sus scrofa
22	110	4.6	456	1	RINI_RAT	P25315 rattus norv
23	109	4.6	388	1	GG2B_MOUSE	P13206 marthasteri
24	107	4.5	480	1	B1N1_HUMAN	P13489 homo sapien
25	106.5	4.5	1403	1	B1N1_HUMAN	Q13075 homo sapien
26	105.5	4.4	1034	1	CARF_HUMAN	Q94P20 homo sapien
27	105.5	4.4	1403	1	B1RE_MOUSE	Q94P20 mus musculus
28	103.5	4.2	1402	1	N14G_MOUSE	Q94W33 mus musculus
29	98.5	4.1	1402	1	B1RG_MOUSE	Q94W33 mus musculus
30	98.5	4.1	1403	1	B1RF_MOUSE	Q94W33 mus musculus
31	96.5	4.1	1172	1	SWP_MOUSE	P41003 schizosacch
32	95	4.0	2778	1	FAP_PROMO	P59324 grosophila
33	94.5	4.0	1447	1	B1RB_MOUSE	Q94W44 mus musculus

RESULT 1	PHY7_HUMAN	STANDARD	PHY7	1062 AA
AC	P59046			
DT	28-FEB-2003	Ref: 41, Created!		
DT	28-FEB-2003	Ref: 41, Last: sequence update!		
DT	15-SEP-2003	Ref: 42, Last: annotation update!		
DE	PRIN1 containing APAF: like protein 7 (March-1)			
GN	PRIN1			
OS	Homo sapiens (Human)			
CC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
CC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo			
CX	NCBI TaxID:9606			
RN	[1]			
RP	SEQUENCE FROM N.A. (ISOPROPY)			
RX	MEDLINE:22162477; PubMed:12019269			
RA	Wang L., Manji G., Grier J.V., Al-Garawi A., Merriam S.,			
RA	Lora C.M., Geddes B.J., Briskin M., Distefano P.S., Bertin J.			
RT	"PRIN1, a novel PYP-N-containing Apaf-like protein that regulates			
RT	activation of NF kappa B and caspase-1 dependent cytokine			
RT	processing"			
ED	J. Biol. Chem. 277:29874-29880(2002)			
RN	[2]			
RP	SEQUENCE FROM N.A. (ISOPROPY) 2 AND 3			
RX	T-SSUE-Lymphoma			
RA	Williams K., Langoff M.W., Harton J.A., Ting J.P.Y.			
RL	Submitted MAY 2002 to the EMBL/GenBank/CCDS databases			
RN	[3]			
RP	SEQUENCE FROM N.A. (ISOPROPY) 1			
RX	T-SSUE=Leukocyte			
RA	MEDLINE:22382577; PubMed:12477912			
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,			
RA	Altschul S.F., Zerkow B., Ruetow K.H., Schaefer C.F., Bhat N.K.,			
RA	Hopkins A.F., Jordan H., Moore T., Max S., Wang J., Hsieh F.,			
RA	Datchenko L., Yarusina K., Parrot A.A., Rubin G.M., Hong L.,			
RA	Slapchenko M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.B.,			
RA	Brownstein M.J., Usdin T.B., Toshiyuki S., Carrin P., Prange C.,			
RA	Rata S.A., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,			
RA	Bosak S.A., McKernan K.J., Malek J.A., Gurarathe P.H.,			
RA	Richards S., Warley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,			
RA	Villalón D.K., Many D.C., Sodergren E.J., Lu X., Gibbs R.A.,			
RA	Fahy J., Heitro E., Kottman M., Yadan A., Rodrigues S., Sanchez A.,			
RA	Whiting R.W., Young A.C., Shevchenko Y., Bouffard G.G.,			
RA	Blakesley A.C., Grimwood J., Schmitz C., Myers R.M.,			
RA	Butterfield Y.S.N., Krzywicki M.J., Skalska U., Smalusz D.E.,			
RA	Schneer A., Schen J.E., Jones S.C.M., Marra M.A.,			
RT	"Generation and initial analysis of more than 15,000 full-length			
RT	human and mouse cDNA sequences"			
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002)			
CC	FUNCTION: May mediate activation of CASP1 via ASC and promote			
CC	activation of NF-kappa B via IKK.			
CC	SUBUNIT: Binds to ASC with its DAPIN domain.			
CC	CELLULAR LOCATION: Cytoplasmic.			
CC	ALTERNATIVE PRODUCTS:			

P34899 pisum sativ
P41738 rattus norv
Q76015 homo sapien
Q22790 homo sapien
Q15413 homo sapien
Q6081C homo sapien
Q19128 brachydanio
P38818 saccharomyc
Q00411 homo sapien
P40026 saccharomyc
Q90224 cyctoisagris
P49454 homo sapien

QY 159 NQKFGA--PRLQPPNPGGLCSLAAGLWCKUFLFEQDLPKHGJGDGV-SAFINX-- 243
 DB 427 LSSAPVAGPRLQ--GDLRLCLRLAREGVLRRAQPAEKELEQELKXGKVQVTLFSKKE 484
 QY 244 --NIFQKDNCE----- 261
 DB 485 LQGVLETVTCFTDQSGFRLAAI SYLELDGGVFRPMAGVGTLLRQDQPHSHLVLT 544
 QY 262 RFGGLLNEERSHLEKSLCWKVSIIKHVLLQWTCRKAQ----- 302
 DB 545 RFLGLASAEKMRDTERHFGCVWSRVKQCBAPWQGGQCGCPVAPEVTEGASGLEDE 604
 QY 303 DQSTLQSS-----JEFFSCYELQESSEFTQQAASHFQVAVVSNIA-SKHEHNVSSPCLK 356
 DB 605 RPEEEEGEERNPYLELLYCLYETQEDAFVRCALCFPELALQVRRCVQVAVLSYCV 664
 QY 357 RCRSAGVLEJ-----XGATYSADTFEARCSCAGATL----- 388
 DB 665 CPGACALELSCRLVAACERKKKSLGRKLCASLGGSSSGGTTKQLPASLJHPLPOMT 724
 QY 389 --LQVLPRTV-----LDPAYSEHIAAACTNPNTLHLYRNALNSPOVKLCCGJRH 442
 DB 725 DPCKHLSLTLSHCKLPDAVCRDLSAURAAFAVTELCMLNRLSPAGLAVLSGLAWPQ 784
 QY 443 CKLQNL 449
 DB 785 CRVQTV 791

RESULT 6
 EVAS MOUSE
 AC Q9:AS2: Q9AC04: FRT, 843 AA.
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 CR PYPIN: containing APAF1-like protein & like
 CS PYPAF5.
 CS Mus musculus (Mouse).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 CX NCBI_TaxId=10092;
 RV [1]
 RF CONCEPTUAL TRANSLATION OF 1-287.
 RA Rinz J.;
 RA Unpublished observations (FEB 2003).
 RN [2]
 RP SEQUENCE OF 288 843 FROM N.A.
 RC TISSUE=Kidney.
 RX MEDLINE=22368257; PubMed=12477332.
 RA Strauberg R.L., Feingold E.A., Grouse L.H., Berger J.B.,
 RA Klausner R.D., Collins F.S., Wagner G., Sherman C.M., Schrier G.C.,
 RA Alfischul S.F., Zeeberg B., Buotow K.H., Schaefer C.F., Abat N.A.,
 RA Hopkins R.F., Jordan H., Moore T., Max A.M., Wang G., Hsieh F.,
 RA Diatchenko L., Varusina K., Farmer A.A., Rubin G.W., Hong J.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Cantoni P., Prange C.,
 RA Raba S.S., Loquellano N.A., Peters G.D., Altman R.D., Mullany S.J.,
 RA Rosak S.A., McEwan P.C., McKernan K.J., Vitek J.A., Gamarathne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hally S.W.,
 RA Villalón D.K., Wuzny D.M., Sodersten E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettner M., Yadan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Soufard G.G.,
 RA Diakostey R.W., Richmond C.W., Green E.D., Jackson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmitz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M., Skalska U., Skallus D.E.,
 RA Scherch A., Schein J.E., Jones S.C.M., Vatta M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences."
 RJ Proc. Natl. Acad. Sci. U.S.A. 99:16693-16693 (2002).
 RE [3].
 RA Albrecht M., Domingues F.S., Schreiber S., Lengauer T.

*Identification of mammalian orthologs associates PYPAF5 with distinct
 functional roles.*
 FEBS Lett. 538:173-177(2003).
 CC 1- FUNCTION: May mediate activation of CASP1 via ASC and promote
 activation of NF-kappaB (By similarity).
 CC 1- SUBUNIT: Binds to ASC with its DAPIN domain (By similarity).
 CC 1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC 1- SIMILARITY: Contains 1 DAPIN domain.
 CC 1- SIMILARITY: Contains 1 NACHT domain.
 CC 1- SIMILARITY: Contains 1 leucine-rich (LRR) repeats
 CC 1- CAUTION: the N-terminus was extended using ESTs and genomic
 sequences, in analogy to ortholog sequences.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).
 CC EMBL: BC013519; AAF13519.1; ALT_INIT.
 CC EMBL: BC031139; AAF11139.1; PYPAF5.
 CC GDB: MG:214:990; PYPAF5.
 CC InterPro: IPR001511; LRR.
 CC InterPro: IPR007391; LRR_RNinh.
 CC Pfam: PF05560; LRR_1.
 CC PROSITE: PS00824; DAPIN; 1.
 CC PROSITE: PS00837; NACHT; 1.
 CC ATP-binding, 11 102 DAPIN.
 CC DOMAIN 164 485 NACHT.
 CC DOMAIN 569 585 ASP/GLU-RICH.
 CC DOMAIN 654 666 POLY-LYS.
 CC REPEAT 434 459 LRR 1.
 CC REPEAT 609 632 LRR 2.
 CC REPEAT 811 834 LRR 3.
 CC NP_BIND 274 151 ATP (POTENTIAL).
 SQ SEQUENCE 843 AA; 94592 MW; 35FE7A766A47D951 CRC64;
 Query Match 28.2%; Score 671.5; DB 1; Length 843;
 Best Local: Similarly 32.5%; Pred No. 1; 6e-48;
 Matches 178; Conservative 86; Mismatches 168; Indels 115; Gaps 15;
 QY 9 SNGDLIFCWPEPSAPLQELIRVPRPLAFITGHPDEKPSHPDQGPWCLQWKEKRPTED 68
 DB 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
 QY 220 SLACAVLQVTPRAWPKRPLAQENRLLFLDGADELPTLPSSSEATPC ---KQLEA 273
 DB 69 69 69 69 69 69 69 69 69 69 69 69 69 69 69 69
 DB 274 TSDRVLSGLSCSELRPGARJLVTRHAATGRLQGLCSPOCAEIRGFSDDKKRYFFK 333
 QY 124 FINAQANVENVYVRCNEPFTMQEFTLVGVVTCLOQQLGGGLRQTSRTTAVYML 183
 DB 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
 DB 334 FDEPKAEPAVEPVKNETLPAQCPVPVCMVCTVQQLQLGRLSLRTSKTTSVVL 393
 QY 184 YLLSLMQK-PGAPRLQPPNPGGLCSLAAG-LWVKILFEQDLRKHGLDGV SAF 243
 DB 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
 DB 394 FTSKXKSNAGTNGPQVQ-GEVTLCLRLAREGILDHKKAFSEEDLEKLRGSGVQTF 451
 QY 241 LNMNIF-----CKDNCE--RSP 256
 DB 452 LMKKEIPGVLTETVYTFIDQSFQFFLAALSYLLEAERTPTGAGGVOKLNSDAELRGH 511
 QY 257 LAUTSRFPGLLNERTSHLEKSLCWKVSPIHKMDLQMIQSKAOSDG ----- 304
 DB 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
 DB 512 LAUTSRFPGLLNERTSHLEKSLCWKVSPIHKMDLQMIQSKAOSDG ----- 571
 QY 305 -----STLQSSLEFFSCLSYIOEERTTQQAASHFQVAVVSNIT-ASKEHNVYS 351
 DB 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
 DB 572 QLEDAHELEEEEDENFG-LESLCYVETQEDAFVRCALCFPELALQVRRCVQVAVLSYCV 630
 QY 357 RCRSAGVLEJ-----XGATYSADTFEARCSCAGATL----- 393

DB 631 NYVCCDDGQCALRVSQGLVAAREKKKKSLUKELKGSOSTKKQPPVSMRLCEITMT 690
QY 394 --- ERTVLG-----PAYSHEIAALATNPNIILSLYNALGSRGVKLCGGKHPN 442
DB 691 TPQHLSVILSHGRIPDAVRDLSFALKAFALKEVLLASRLTNIGFLCEGLAMPK 750
QY 443 CKLQNLK 449
DB 751 CQVKTLR 757

RESULT 7
ID PYA5-RAT STANDARD; PRT; 354 AA.
AC C63035;
DT 28-FEB-2003 (Rel. 42, Last sequence update)
DT 28-FEB-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE PVR-N-containing APAF1 like protein 5 like (Anopheles gambiae) 11/vasopressin
DE receptor;
DN PYA5 OF AVR.
OS Rattus norvegicus (Rat);
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus;
CX NCBI TaxID:10116;
RN 1;
RP CONCEPTUAL TRANSLATION OF 1-343.
RA Ruz-Zapata N., Lopez L.V., Herrera V.M.M.;
RX MEDLINE=96371640; PubMed=7453366;
SA Ruz-Zapata N., Akimoto K., Herrera V.M.M.;
RT Identification of a novel dual androgen/androgen 11/vasopressin receptor on
the basis of molecular recognition theory;
RL Rat. Med. 1:1074-1082(1995);
RN 1;
RP VARIANTS SALT-SENSITIVE HYPERTENSION SEP 492 AND ARG 536.
EX VHELINE219K185; PubMed11984503;
RA Ruz-Zapata N., Lopez L.V., Herrera V.M.M.;
RT The dual And1/AVP receptor gene N11WS/11AP variant exhibits
sodium-induced dysfunction and coexpressed with salt sensitive
hypertension in the Dahl salt-sensitive hypertensive rat model;
RL Mol. Med. 8:224-227(2002);
RN 1;
RP IDENTIFICATION OF XAMMALAN (XAM) FOR THE PYA5;
RA Albrecht M., Domingues F.S., Schreiber S., Lehmann T.;
RT Identification of mammalian orthologs associated with PYA5 with distinct
functional roles;
RL FEBS Lett. 538:173-177(2003);
CC 1. FUNCTION: May mediate activation of CASP1 via ASC and promote
activation of NF-kappa-B (by similarity). Androgenin 11 and
vasopressin binding protein. May stimulate cAMP accumulation.
CC 2. SUBUNIT: Binds to ASC with its CAPIN domain (by similarity).
CC 3. SUBCELLULAR LOCATION: Cytoplasmic (by similarity).
CC 4. TISSUE SPECIFICITY: Highly expressed in kidney. Detected at low
levels in all tissues tested.
CC 5. DISEASE: Defects in PYA5 may be a cause of salt sensitive
hypertension.
CC 6. SIMILARITY: Contains 1 CAPIN domain.
CC 7. SIMILARITY: Contains 1 NACTH domain.
CC 8. SIMILARITY: Contains 4 leucine-rich (LRR) repeats.
CC 9. CAUTION: The N-terminus was extended using ESTs and genomic
sequences, in analogy to ortholog sequences.
CC 10. This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

CC or send an email to license@sib-sib.ch.
DB EMBL: M85183; AAA03623.1; ALT_INIT.
DR InterPro: IPR001411; LRR_RNinh.
DR Pfam: PF00560; LRR_2.
DR Pfam: PF02758; PFAD-DAPIN; 1.
DR PROSITE: PS50624; DAPIN; 1.
DR PROSITE: PS50837; NACTH; 1.
KW ATP-binding; Leucine-rich repeat; Repeat.
FT DOMAIN 1 133 DAPIN.
FT DOMAIN 169 484 NACTH.
FT DOMAIN 549 604 ASP/GLY RICH.
FT DOMAIN 659 662 POLY-LYS.
FT REPEAT 433 458 LRR 1.
FT REPEAT 613 633 LRR 2.
FT REPEAT 722 745 LRR 3.
FT REPEAT 812 836 LRR 4.
FT NP_BIND 174 181 ATP (POTENTIAL).
FT VARIANT 492 492 N -> S IN SALT-SENSITIVE HYPERTENSION.
FT VARIANT 536 536 C -> R IN SALT-SENSITIVE HYPERTENSION.
FT MUTAGEN 766 766 E -> K ABOLISHES ANDROGENIN 11 BINDING.
SQ SEQUENCE 354 AA; 95293 MW; DHRFC922D7B3734 CRC64.

Query Match: 27.6%; Score 695; DB 1; Length 854;
Best local similarity: 32.1%; Pred. No. 3.9e 47;
Matches 170; Conservative 92; Mismatches 124; Gaps 13;

QY 9 SMOGLFSPSPSAPQGLIRVRLFLDGGDFELKPSFHDGPGACLOWEKRPTEL 68
DB 220 SNAFLDECPORAPVRILAQPHRLFLDGADELPTLAPEATPC----PCPEA 273
QY 69 -- LENSIRPKKLIFELSSLTTPFALEXHRLLEPRHVELGFSBAERKEPYKY 123
DB 274 TSULRVLSGLSGLSPSASLLVTSRNATLGLQGLSPCAEVGRGFSKDKKKYRFE 333
QY 124 ENMAEAGQVNVVRDNEPFTMVFVLVWVVTCTLOOGLGGSLTQRTTAVNL 183
DB 334 FMDERKAEAVRFVRENETYALCFVFPVCMVTCTVLQQKELGRDLSTRTTTSVY 393
QY 184 YLLSLVQPKFGAPRQPPNCRGLCSLAAGSLWCKILFEBODURKKHGLGDEYSA 239
DB 394 FITSKLKGASTNPRVQ--GELRMCHLAAREILKHQGFSEKLERLKQSSQCTMFL 451
QY 240 --- FUNKN--- FCKDINCE RSFL 257
DB 452 SKRELEPRIVVYTCFIDQSPQFELALNSYLDAGAPGNSAGSVQMLRSALRPHL 511
QY 458 AATSPRELKQLNEFESHLEKSKKKVAPHIKWLLQMGSKAQSGSTL----- 307
DB 512 AATRTPLGLNSTERHICGNHFGVVPGRVQDRLRWQGSQPKVATVGAKKDELKC 571
QY 308 --- QGSLFFESCLYEIQEEFQQALSHFQVIWSTASKMEHYVS 351
DB 572 EEAHREHEEHEEERKAFGLGLYCLVETCEDFVRQALSSLPENVLERVTRVDLEV 631
QY 352 SPGLKRCRAQVLEHGNATYSAGGEPAR-----CSAGAH---TLVGLPERIV 397
DB 632 SYGVQGGDQALRVSQGLVAAREKKKKKFXNRLKGSQSTGKQPPASLRPLCEAY 691
QY 398 --- LLDAYSEHMAALCTNPNILDSYRNALGSRGVKLCGGKHPN 442
DB 692 TQOGLSLTSLSHCKLPDAVRCDSFAJKVAPSLBELGLQNLRLTEAGIRLSQGLANK 751
QY 443 CKLQNLK 449
DB 751 CQVKTLR 758

RESULT 8
NAL4 HUMAN
ID NAL4 HUMAN STANDARD; PRT; 994 AA.
AC Q96MN2; Q96AV6;

DT 28-FEB-2003 (Ref. 41, last sequence update);
 DT 15-SEP-2003 (Ref. 42, last annotation update);
 DE NACHT¹, LRR- and PYD-containing protein 4 (PAAD) and NACHT-containing
 DE protein 2 (PYRIN-containing APAF-like protein 4) (Pibonclease
 DE inhibitor 2).
 GN NALP4 OR PYPAF4 OR PAN2 OR RXH2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catartini; Hominoidea; Homo.
 OX NCBI_TaxID=9606;
 RN [1]_TaxID=9606;
 RP SEQUENCE FROM N.A. (ISOFORM 1);
 RA Xatlinopoulos F., Ischoff J.;
 RT "NALP4, a novel member of the PYD, NACHT, and LRR family";
 RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM 1);
 RA Florentino L., Reed J.C.;
 RT "NALP4, a novel PAAD-containing protein";
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORM 1);
 RA Retin J.;
 RT "PYPAF4, a novel PYRIN-containing APAF-like protein";
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A. (ISOFORM 1);
 RA Miyamoto T.;
 RT Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A. (ISOFORM 1);
 RA Niimiya K., Wagatsuma M., Kanda K., Kondo H., Yokoi T., Kodaira H.,
 RA Fukuoka T., Takahashi M., Kikkawa E., Umura Y., Ake K., Kimihata K.,
 RA Katsuta N., Sato K., Takikawa M., Yamazaki M., Sugiyama T., Irie R.,
 RA Cusick T., Sato H., Wakamatsu A., Iwata S., Yamashita M., Isono Y.,
 RA Kawai H., Saito K., Nishikawa T., Komura K., Yamashita H.,
 RA Matsuo K., Nakamura Y., Sekine N., Kikuchi H., Murakawa K.,
 RA Kurehara K., Takahashi Fujii A., Oshima A., Sugiyama A., Kawakami B.,
 RA Suzuki Y., Sugata S., Nagahata K., Masuda Y., Nishii K., Isonai T.;
 RT "NERO human cDNA sequencing project";
 RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SEQUENCE OF 436-994 FROM N.A. (ISOFORM 1);
 RC TISSUEmaplcent;
 RX MEDLINE=22386257; PubMed=12477930;
 RA Strassberg R.L., Feingold E.A., Groble G.H., Berge G.J.,
 RA Klausner R.D., Collins F.S., Wagner T., Schmedt C.M., Schuler G.D.,
 RA Altshuler S.F., Zeeberg A., Buetow K.H., Schaefer C.F., Bat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.L., Wang C., Hsieh P.,
 RA Diatchenko L., Marusina K., Farber A.A., Rubin G.M., Hong F.,
 RA Shapiro M., Soares M.B., Bonaldo X.F., Casavant T.J., Schetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Gaidin P., Prange C.,
 RA Raba S.S., Lequellano N.A., Peters G., Anderson E.P., Mullany S.J.,
 RA Bonak S.A., McEwan P.J., McKernan K.G., Valler J.A., Gharatne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A., Guy L., Bulik S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E., Lander E., Gibbs R.A.,
 RA Fahy J., Helton E., Kottman M., Madan A., Patrino S., Sanchez A.,
 RA Whitting R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Meyer R.V.,
 RA Rutterfield Y.S.N., Krzywicki M., Skalski J., Scallan D.E.,
 RA Schreier A., Schein J.E., Jones S.M., Martin M.A.,
 RA "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:14876-14881 (2002).
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event:Alternative splicing; Named isoforms;
 CC isoform=1;
 CC Name=1;
 CC IsoId=Q96MN2-1; Sequence=Displayed;
 CC Name=2;

CC IsoId=Q96MN2-2; Sequence=VSP_003917;
 CC Name=3;
 CC IsoId=Q96MN2-3; Sequence=VSP_003916;
 CC -!- SIMILARITY: Contains 1 DAPIN domain;
 CC -!- SIMILARITY: Contains 1 NACHT domain;
 CC -!- SIMILARITY: Contains a leucine-rich (LRR) repeats
 CC extend the sequence.
 CC This SWISS-PROT entry is copyrighted, it is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib.ch).
 CC -----
 DR EMBL; AF42488; AAL35293.1;
 DR EMBL; AF42792; AAL6396.1;
 DR EMBL; AF47947; AAL7104.1;
 DR EMBL; AF48206; AAL8672.1; ALT_NIT.
 DR EMBL; AKC5668; BAB2254.1;
 DR EMBL; RC01643; AAL1644.1;
 DR InterPro: IPR020701; LRR_RNase.
 DR InterPro: IPR037111; NACHT_NTRase.
 DR InterPro: IPR046233; PAAD_DAPIN_dom.
 DR Pfam: PF02758; PAAD_DAPIN_1.
 DR PROSITE: PS0824; DAPIN; 1.
 DR PROSITE: PS0837; NACHT; 1.
 KW Repeat; Leucine-rich repeat; Alternative splicing.
 FT DOMAIN 1 94
 FT DOMAIN 149 472
 FT REPEAT 637 663
 FT REPEAT 668 721
 FT REPEAT 722 745
 FT REPEAT 750 777
 FT REPEAT 780 833
 FT REPEAT 836 866
 FT REPEAT 868 943
 FT REPEAT 949 972
 FT VARSFLIC 1 93
 FT MAASFFSDFGLMWLEELKSEFKFKFKELKQMTLOLELKO
 FT LFWKVKASREKANIILKIVEEQAWNITILIFQWYPRK
 FT DLMKAVYREPT -> MQECLTLWVSPALIDS IIR
 FT isoform 3)
 FT /FTID=VSP_003916.
 FT Missing in isoform 2).
 FT /FTID=VSP_003917.
 FT L -> P (IN REF. 4).
 FT Y -> H (IN REF. 4).
 FT I -> T (IN REF. 4).
 FT W -> R (IN REF. 4).
 FT I -> V (IN REF. 4).
 FT Y -> C (IN REF. 4).
 FT L -> P (IN REF. 4).
 FT SEQUENCE 994 AA; 113444 MW; 1CECF68B1839C9 CRC64;
 Query Match 27.5%; Score 652.5; DB 1; Length 994;
 Best Local Similarity 31.1%; Pred. No. 7.8e-47;
 Matches 153; Conservative 98; Mismatches 170; Indels 7; Gaps 7;
 QY 9 SMOGLFSCWPEPSAPLOELRVPERLLFIIDGFDLPKSPFDPOGPPWCLCWEKRPTEL 68
 DB 200 SVALDISREWDPAPIITEIVSQPERLLFVIDSFEELQGSNEPDSLDGLMEKRPVQV 259
 QY 69 LNSLIRKKLFLPESLDTTTRPTALEKHRLLEHPRHVEILGFSEAEKKEYFYKFNAAE 128
 DB 260 LLSLLLPKKMGFEASLJAIRVCPKELRDQVTISEIYQPRGFNEDRLVVFCCFFDKP 319
 QY 129 QACQVENVVNEDEPLTFVCPVPLVGVWVCTTLCQQLBGGLLRQTSRTTITVYKYLSL 188
 DB 100 RAMEAFNLVSESEQLFSQCFILWLLCTALKQRMQKGLALTCQSTTSVYSFVNL 379
 QY 189 YGPKGKA FRLQPPFNGRGLCSHAADGLWPKLFLFEQDLRKHGLDGDVSAFINMNF 246


```

Co 360 FTPE-GREGPTQCHQKALCSAAEWK...PEP...PRNVVIAIPALISKIL 436
QY 347 QXDINCEPSFAS... 259
DB 439 LKYGERSSVYFVJVCIGERTFAALFYLRSHLHHHIAVGVVELVANPERARRAIFIP 498
QY 260 TSKRQFGLINETHSLKSLKNSKVSSEHKMKMLLOWCSKAQSDGSLDQGSLEFFSCSLY 319
DB 499 LGCFUTSLNKKSEKIDAFKQFQSGSEIKQVHCQKSLSEKIDPQPCQVDSNAIFVCLF 559
QY 320 EIQSEETFCQALSHFQV... WQSTIAKSHHWYVSELEFQKSAVLUHUYGATYSADGE 376
DB 559 ENQDPAFVKQAVNLQGANSEHILNV... DUVSAFALVVSSELEKICF SVQNVFKK 612
QY 377 DRACSCAGATLLVQERTVLLVAYSEHAAALTNINLEHLYNNAQSGKVLQCO 436
DB 613 EDEHSTSDYS...W...HHGVSVLTSSRLKELQVQSTSESTFVWCV 660
QY 437 GRHPNCKIQKL 448
DB 561 QRHPSCRLQKL 672

RESULT 3
NAME HUMAN
ID NAME HUMAN STANDARD, FREE, 1740 AA
AC QNKK22, QNKKV5, QNKK56, QNKKV9, QNKKV1
DT 14-OCT-2001 (rel. 42, last sequence update)
DE 16-OCT-2001 (rel. 42, last annotation update)
FE NACHT-1, LRK- and PYD-containing protein 2, Nucleotide binding site
PE Protein 1, PYRIN-containing APT-1-like protein 2
GN NACHT-1, CS NEST, CS PYRAF2
OS Homo sapiens (human)
CC Eukaryotic, Metazoa; Chordata, Vertebrata; Euteleostomi;
CC Vertebrata; Euteleostomi; Primates; Catarrhini; Hominoidea; Homin
CC NCBI TaxID:9606
RN 1
RP SEQUENCE FROM N.A. (ISOGEM 1)
RX NELLINE21169419, PubMed:11276169
RA Heitlin J., Gasteigano P.S.
RT "The PYRIN domain: a novel motif found in apoptosis and inflammation
RT proteins."
RC Cell Death Differ. 7:1173-1184 (2000)
RN 121
RP SEQUENCE FROM N.A. (ISOGEM 1)
RX NELLINE21148093, PubMed:11276169
RA Kattinon F., Heilmann K., Eschopp G.
RT "The PYRIN domain: a possible member of the CARD family
RT implicated in apoptosis and inflammation."
RC Curr Biol. 11:R118-R120 (2001)
RN 13
RP SEQUENCE FROM N.A. (ISOGEM 1)
RX NELLINE21169427, PubMed:12102669
RA Wang L., Wang L.G., Grenier J.V., Al-Gharab A., Vothman S.,
RA Zota J.M., Geddes B.J., Priskin V., Kishimoto A.S., Berlin J.
RT "PYRAF2, a novel PYRIN-containing Apat-1-like protein that regulates
RT activation of NF-kappa B and caspase-1-dependent cytokine
RT processing."
RC J. Biol. Chem. 277:29574-29584 (2002)
RN 14
RP SEQUENCE FROM N.A. (ISOGEM 1)
RX NELLINE21169427, PubMed:12102669
RA Kikarabe K., Kuragai A., Iikura S., Yokoyama M., Takahashi H., Ota T.,
RA Suzuki Y., Ohyasui Y., Nishi T., Shibata T., Iizuka T.,
RA Nakamura Y., Isogai T., Sugano S.
RT "NEMO human cDNA sequencing project."
RC Submitted (FEB-2002) to the EMBL/GenBank/CCDB databases
RN 15
RP SEQUENCE FROM N.A. (ISOGEM 1)
RX NELLINE21169427, PubMed:12102669
RA Kikarabe K., Kuragai A., Iikura S., Yokoyama M., Takahashi H., Ota T.,
RA Suzuki Y., Ohyasui Y., Nishi T., Shibata T., Iizuka T.,
RA Nakamura Y., Isogai T., Sugano S.
RT "NEMO human cDNA sequencing project."
RC Submitted (FEB-2002) to the EMBL/GenBank/CCDB databases
RN 16
RP SEQUENCE FROM N.A. (ISOGEM 1)
RX NELLINE21169427, PubMed:12102669
RA Kikarabe K., Kuragai A., Iikura S., Yokoyama M., Takahashi H., Ota T.,
RA Suzuki Y., Ohyasui Y., Nishi T., Shibata T., Iizuka T.,
RA Nakamura Y., Isogai T., Sugano S.
RT "NEMO human cDNA sequencing project."
RC Submitted (FEB-2002) to the EMBL/GenBank/CCDB databases

```

```

RA Strausberg R.L., Feingold R.A., Grouse L.H., Gorge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.C.,
RA Altschul S.F., Zuercher B., Suetow K.H., Schaefer S.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Vax S., Wang J., Hsieh J.,
RA Datschenko W., Marusta K., Parmer A.A., Rubin G.M., Hong J.,
RA Stauffer M., Soares M.B., Ronald M.F., Casavant T.L., Schetz T.E.,
RA Brownstein M.C., Udell J.B., Tschiszyki S., Carninci P., Prange C.,
RA Raha S.S., Beckelund N.A., Fetsch G.J., Abramson R.D., Mullen S.C.,
RA Bosak S., Melnik B.J., McKernan K.J., Valek J.A., Gagarin P.H.,
RA Richards S., Ashley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.M., Mizny D.M., Sodergren E., Xu X., Gibbs P.A.,
RA Welling J., Heitlin J., Kottmann M., Madan A., Rodrigues S., Sanchez A.,
RA Mahtani M., Nadeau A., Young A.C., Shevchenko Y., Boulard G.G.,
RA Blakesley R.W., Townsend J.W., Green E., Dickson M.C.,
RA Rodriguez A.C., Greenwood J., Schuttz J., Myers R.M.,
RA Butterfield Y.S.R., Krzywicki M., Skalska U., Smalusz D.E.,
RA Scherch A., Schen J.E., Jones S.M., Mitra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences."
RT Proc. Natl. Acad. Sci. U.S.A. 99:16833-16837 (2002).
CC 1- FUNCTION: May be implicated in apoptosis (By similarity).
CC 1- CORRELATION: Birds, ATP (By similarity).
CC 1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC 1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms:2;
CC Name=1;
CC iso1 (QNKK22) 1; Sequence=Displayed;
CC Name=2;
CC iso2 (QNKKV5) 2; Sequence=VSP_055522;
CC 1- SIMILARITY: Contains 1 PYRIN domain.
CC 1- SIMILARITY: Contains 1 NACHT domain.
CC 1- SIMILARITY: Contains 1 Leucine-rich (LR) repeats.
CC This SwissProt entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sb.ch).
CC
CC EVBL: AB298549, AAC18263, 1; AUT INIT.
CC EVBL: AB310186, AAC3289, 1;
CC EVBL: AB444761, AAC4962, 1;
CC EVBL: AK015513, BAA1223, 1;
CC EVBL: AK030623, BAA16294, 1;
CC EVBL: AK042864, BAA1173, 1; AUT INIT
CC EVBL: BC016941, BAA1504, 1;
CC EVBL: BC016941, BAA1649, 1;
CC HSSP: F00755, 4RHE
CC InterPro: IPR027091, LPS_RRM1
CC InterPro: IPR027111, NACHT_NTPase
CC InterPro: IPR049224, PAAL_DAPIN_Rom.
CC PROSITE, PS00474, DAPIN
CC PROSITE, PS00477, NACHT
CC Apoptosis: AIP-binding, leucine rich repeat; Repeat;
CC Alternative splicing:
CC DOMAIN 1 94 DAFIN
CC DOMAIN 226 526 NACHT
CC REPEAT 407 492 LRR 1
CC REPEAT 402 445 LRR 2
CC REPEAT 414 477 LRR 3
CC REPEAT 410 442 LRR 4
CC REPEAT 416 451 LRR 5
CC REPEAT 402 442 LRR 6
CC REPEAT 404 442 LRR 7
CC REPEAT 403 442 LRR 8
CC REPEAT 403 442 LRR 9
CC REPEAT 403 442 LRR 10
CC NP_5190 POLY-AD
CC DOMAIN 1 94 AIP-INTERACTING
CC VASISL1 1 154
CC CONFIDENT 1 X V IN REPEAT 1

```

FT CONFLICT 15 35 10 -> P (IN REF. 11)
 FT CONFLICT 304 364 10 -> S (IN REF. 4); BAB152931
 FT CONFLICT 364 364 P -> K (IN REF. 1)
 FT CONFLICT 980 980 MISSING (IN REF. 1)
 FT CONFLICT 1052 1052 A -> E (IN REF. 3)
 SQ SEQUENCE 1362 AA: 120514 XM, 4DBH0F6E9240CA; CRC644

Query Match 23.7%; Score 544; DB 1; Length 1062;
 Best Local Similarity 25.9%; Pred. No. 2.4e-39;
 Matches 171; Conservative 83; Xismatches 175; Indels 130; Gaps 16;

QY 8 CSMQLIFSCFESAPLCHLRVPERLEFIDGHLKINSEHNQHWLCKEKEKPE 67
 DB 256 CSTAGLVPRDPPKQDDPHPLAARAKLFFVDSDFCEIGAPALCIEDICGEPKZKPEVP 375
 QY 68 ILLSLIKKKLPELSLITTEPALEKJHKLSEHRIVEILGSEAEKREYFYKVENNA 127
 DB 316 VLLGLSLKRVMLPKAAALIVTPRPAQRDLRLAREPYIRVEGFLIEDRRAYFLIRHGE 375
 QY 128 EOAGOVNYPVNEPFFYCHVPLVGVVCTCQCGLEGGLGRLQTSRTTAVVYLLS 187
 DB 376 PQAVRAFFLMESNAALFGLGAPAVGVCTCTKQWKGEDPVPETGLTGLFRLCS 435
 QY 188 LMCPKPAFQPPNQSLCSLAAGLWKKLIFEDCTLEKRGLEDVYSANWNIQC 247
 DB 436 RF PQQAQUGA LATELLAAQUGLNAQTVLHREDLFLVQVSDLRHFLGGILR 490
 QY 248 KD INCGSFALTSKFL --- --- --- --- --- --- --- --- --- 264
 DB 421 QGVSKGC YSFHLSFCFLTALPYTLKKEEHRDRTQWQIGVQKLSGVRLRNP 549
 QY 265 --- --- --- --- --- --- --- --- --- --- --- --- --- 314
 DB 530 LQAGVYSGLANEKRAKELEATQFQKSTFKQELSLDITKAGSSTVQLQ --- --- --- 575
 QY 315 FSLVETCEEFIOALSHQVIVVSKASMKHWLSEFLKESQAQVHL --- --- --- 366
 DB 606 LCCVYESGSEVREVMQAQNEI SLEHMAVTVVSEFVHCHNLQKMSQVKEKELP 663
 QY 367 --- --- --- --- --- --- --- --- --- --- --- --- --- 369
 DB 664 ENVTASEDAEVSQDDCHMPEWTLKSTENSKMLKALINQSLASVRLDCEQI 723
 QY 370 --- --- --- --- --- --- --- --- --- --- --- --- --- 378
 DB 724 ASOTCHQGVVFNKSTADAHNE LAKHREYVYVTEFNDQWLMQEPALCEVDRHPEC 783
 QY 379 --- --- --- --- --- --- --- --- --- --- --- --- --- 393
 DB 784 NURYGLVUSGATTCQNAQSLALVNLSTVQWQNLQDQAKILVTLHKKFELQ 843
 QY 394 --- --- --- --- --- --- --- --- --- --- --- --- --- 448
 DB 844 RSLNCHLTANCKDLAAVLVSEKTHQTAQVSLHTVWVFCGLHRYPECKLQI 902

RESULT 10
 PYA3 HUMAN
 ID PYA3 HUMAN STANDARD; PRT; 480 AA;
 AC Q8WX94;
 DT 28-FEB-2003 (Rel. 41, Created;
 DT 28-FEB-2003 (Rel. 41, Last sequence update;
 DE 28-FEB-2003 (Rel. 41, Last annotation update;
 DE PYRIN-containing APAF1-like protein 1;
 GN PYPAF1;
 OS Homo sapiens (Human);
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo;
 OX MCB; TaxID=9606;
 RN 11;
 RP SEQUENCE FROM N.A.
 FX MEDLINE=22162427; PubMed=12319269;
 RA Wang L., Manji S.A., Grenier J.M., Al-Gatwari A., Mertin S.,

RA Lora J.M., Geddes B.C., Atskin X., DiStefano P.S., Bertin J.,
 RT "PYPAF7, a novel PYRIN-containing Apaf1-like protein that regulates
 RT activation of NF-kappa B and caspase 1-dependent cytokine
 RT processing";
 RJ Biol. Chem. 277 23874 23880 (2002)
 CC 1- SIMILARITY: Contains 1 DAPIN domain.
 CC 1- SIMILARITY: Contains 1 NACHT domain.
 CC 1- SIMILARITY: Contains 9 leucine-rich (LRR) repeats.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib.ch).

EMBL AF464765; AAL63961.1;
 InterPro: IPR007591; LRR_RNTH;
 DR InterPro: IPR007111; NACHT_NTPase;
 DR InterPro: IPR004020; PAAD_DAPIN_3nt;
 DR Pfam: PF02756; PAAD_DAPIN_3nt;
 DR PROSITE: PS50424; DAPIN_1;
 DR PROSITE: PS50437; NACHT_1;
 KA ATP-binding; Leucine-rich repeat; Repeat;
 FT DOMAIN 1 93 DAPIN;
 FT REPEAT 172 491 NACHT;
 FT REPEAT 514 638 LRR 1;
 FT REPEAT 674 697 LRR 2;
 FT REPEAT 760 784 LRR 3;
 FT REPEAT 798 820 LRR 4;
 FT REPEAT 817 847 LRR 5;
 FT REPEAT 845 868 LRR 6;
 FT REPEAT 874 897 LRR 7;
 FT REPEAT 903 928 LRR 8;
 FT REPEAT 933 957 LRR 9;
 FT NE_BINT 194 195 ATP (POTENTIAL);
 SQ SEQUENCE 980 AA: 111806 MW: 822AF2F419303D CRC64;

Query Match 23.7%; Score 561.5; DB 1; Length 980;
 Best Local Similarity 25.8%; Pred. No. 2.3e-39;
 Matches 169; Conservative 83; Xismatches 179; Indels 24; Gaps 14;

QY 8 CSMQLIFSCFESAPLCHLRVPERLEFIDGHLKINSEHNQHWLCKEKEKPE 63
 DB 248 CSPALLKQWELQGLTSLTACQPIFYVDSGLDLKV---PPGALCQICGDWEKK 276
 QY 64 RPTETLLKSLPKKLPKLSLITTEPALEKIHLLSHPRHVEILGSEAEKREYFYKY 123
 DB 277 KPVVLLSLKPKVPSAALIVTPRPAQRDLRLAREPYIRVEGFLIEDRRAYFLRH 336
 QY 124 ENAQAQGVNYPVNEPPLFMCTVPLVGVVCTCQCGLEGGLLQTSRTTAVVYL 183
 DB 337 FDEQDAVRAPELMSNAALFGLGAPAVGVCTCTKQWKGEDPVPETGLTGLF 196
 QY 164 VLLSLYCHKPKAPRPGCTFENQGLCLANDGLKNGKILFECDLKHGLDGSVAFINM 243
 DB 397 FLCSMF PQQAQUGA LATELLAAQUGLNAQTVLHREDLFLVQVSDLRHFLGG 451
 QY 244 NIFQKD---INCGSFAL---INCGSFAL---INCGSFAL---INCGSFAL--- 259
 DB 452 DLRQGVSKGCYSEFHLSPQFATLFLYAEKEGEDRDGHANDIGPVQKLSGEERLK 511
 QY 260 --- --- --- --- --- --- --- --- --- --- --- --- --- 313
 DB 513 RFLQVGHFVGLANEKRAKELEATQFQKSTFKQELSLDITKAGSSTVQLQ 570
 QY 314 FSTVYVEGEHFIQALSHQVIVVSKASMKHWLSEFLKESQAQVHL 369
 DB 571 VLGDFESQBELARVVAVAPNEISIHLTNTSEVMHCSFLKHQCDLCKLSQVAKGVF 629
 QY 393 --- --- --- --- --- --- --- --- --- --- --- --- --- 474

"A mouse gene encoding an oocyte antigen associated with autoimmune premature ovarian failure";
 Endocrinology 140:3720-3726(1999).

SEQUENCE FROM N.A. AND VARIANTS ASP-7; LYS 42; LYS-87; PRO-256;
 GLN-977 AND 1003-ASN ASN-1004.

GLN-977 AND 1003-ASN ASN-1004;
 STRAIN=29/SV;

MEJLINE=20222768; PubMed=12754191;
 TONG Z.-B., Nelson L.M., Dean C.

"Water encodes a maternal protein in mice with a leucine-rich repeat domain homologous to porcine ribonuclease inhibitor";

Nat. Genet. 11:281-287(2000).

FUNCTION.

MEJLINE=20517328; PubMed=1062459;
 TONG Z.-B., Gold L., Pfeiffer K.E., Dorward H., Lee E., Hoady C.A.,

Dean J., Nelson L.M.;
 "Water, a maternal effect gene required for early embryonic development in mice";

Nat. Genet. 26:267-268(2000).

FUNCTION: Necessary for embryonic development beyond the 2-cell stage.

CELLULAR LOCATION: Cytoplasmic.

TISSUE SPECIFICITY: Oocyte-specific.

SIMILARITY: Contains 1 NACT domain.

SIMILARITY: Contains 13 leucine rich (LRR) repeats.

This SWISS-PROT entry is copyright © 1999 and is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed (stage by and for commercial entities requires a license agreement (see <http://www.ebi.ac.uk/ebis/announce/> or send an email to license@ebi.ac.uk).

EMBL: AF074028; AAD5762.1; MGD: YGL1345-93; Mater.
 EMBL: AF143573; AAF64393.1; JOINED.
 EMBL: AF143559; AAF64393.1; JOINED.
 EMBL: AF143560; AAF64393.1; JOINED.
 EMBL: AF143561; AAF64393.1; JOINED.
 EMBL: AF143562; AAF64393.1; JOINED.
 EMBL: AF143563; AAF64393.1; JOINED.
 EMBL: AF143564; AAF64393.1; JOINED.
 EMBL: AF143565; AAF64393.1; JOINED.
 EMBL: AF143566; AAF64393.1; JOINED.
 EMBL: AF143567; AAF64393.1; JOINED.
 EMBL: AF143568; AAF64393.1; JOINED.
 EMBL: AF143569; AAF64393.1; JOINED.
 EMBL: AF143570; AAF64393.1; JOINED.
 EMBL: AF143571; AAF64393.1; JOINED.
 EMBL: AF143572; AAF64393.1; JOINED.
 PIR: A59000; A59000.

RSSP: P10775; 2BNH.

MGD: YGL1345-93; Mater.

GO: GO:0005829; Cytoosol; IDA.

GO: GO:0009887; P:cytoogenesis; IPI.

InterPro: IPR001611; LRR.

InterPro: IPR007391; LRR_RNinh.

InterPro: IPR007111; NACT_NTPase.

Pfam: PF00560; LRR_3.

PROSITE: PSS0837; NACT; 1.

ATP-binding: Leucine-rich repeat: Repeat: Poly-nib.sm.

DOXAIN 1 128 5 X APERXINATE TALEM REPEATS.

REPEAT 1 23 1 (INCOMPLETE).

REPEAT 24 49 2.

REPEAT 50 75 3.

REPEAT 76 102 4.

REPEAT 103 128 5.

REPEAT 191 513 NACT.

REPEAT 540 664 IPI.

REPEAT 691 714 LRR 2.

REPEAT 737 762 LRR 3.

REPEAT 776 799 LRR 4.

FT REPEAT RC4 832 LRR 5.
 FT REPEAT 834 856 LRR 6.
 FT REPEAT 861 884 LRR 7.
 FT REPEAT 890 913 LRR 8.
 FT REPEAT 918 941 LRR 9.
 FT REPEAT 950 974 LRR 10.
 FT REPEAT 978 1003 LRR 11.
 FT REPEAT 1025 1027 LRR 12.
 FT REPEAT 1032 1059 LRR 13.
 FT NP_BIND 197 204 ATP_POTENTIAL.
 FT VARIANT 7 7 E -> D (IN STRAIN 129/SV).
 FT VARIANT 42 42 G -> L (IN STRAIN 129/SV).
 FT VARIANT 87 87 R -> K (IN STRAIN 129/SV).
 FT VARIANT 256 256 D -> P (IN STRAIN 129/SV).
 FT VARIANT 977 977 H -> Q (IN STRAIN 129/SV).
 FT VARIANT 1003 1004 SS -> NN (IN STRAIN 129/SV).
 SQ SEQUENCE 1111 AA; 125502 MW; 4016A5D6VAICCP4 CRC64;

Query Match 22.4%; Score 545; DB 1; Length 1111;

Best local similarity 29.1%; Pred. No. 6.9e-37;

Matches 143; Conservative 85; Mismatches 307; Indels 56; Gaps 6;

QY 9 SMQLIFSCWPPSPAHQLQELIVPERLFLICGFDELKPSFDPOGPMCLCHEKRPTEL 69
 DB 241 SLAQLIAKCPDSNOLIVTKIMSQPELLFVLDGLDMSVLQHDQMTLRCKKQSQPIVI 300
 QY 69 LNSLIRKKLDELSELITTRPTALEKLRLLERPHVEILGFSEAEKKEYKVFHNAE 128
 DB 301 LKYSLLRKALQCSFELITTRNTGLEKLSMVSYPYLVEGLSASRRSCLVIENTSNES 160
 QY 129 QAGQVENVYRDEPFTKCFVFLVQWVTCIQQLQEGGLLRQTSRTTAYVMIYL 185
 DB 361 LRIGVFHSLIENHQIFDQCAPSVCSLQCEALQKLGKRCCTLPCQNLGLVATLVEHQ 420
 QY 186 LSIQMKKAPRLQPFNQSGLSLAAGLKNQKLFEEQDLKKGGLGEVSAFLNNNI 245
 DB 421 LTKRPSQALSQERQITLVGLQMAKAGVMTMRSVFYDDDLKNYSLKESLALFENK 480
 QY 246 FQKQNCERSFLAL-----FLAL----- 259
 DB 481 LLOVHNSECCYVFVSHLSIQDFEALYVLEGLSENNQHFCEINQPSIYVKRTDTRL 540
 QY 260 LSKPLQGLNEFTSRSLKSLQKVKVSPHLYKYLQWQSKAQSGSTLQGSLEFFSC 317
 DB 541 LQVKEPRLGLMKDILKTLVFLFVFTVPTVEKQLHWVSLIAQCVNGSTSPMDTLDAFVC 600
 QY 318 LVEIYGEELIQALSHFOVLVSNKASKWEHVVSFCIKRCSAQVLHLYGATYSAGGED 377
 DB 601 LFESQDERFWAGALKRFPQEWILLNQKMKVSSYCYLKHQCNLKAIRVETFDLLSVDMT 659
 QY 378 KARGSAGHTILVQIPERTVLLDAYSEHLAAALCTNPDLIELSLYRNALSGRGVLLCOG 437
 DB 660 LRLCP---VVIVQETQCKPFLMEWGNFCSVLGSLRNKELDLGDSILSORAKMLCLE 715
 QY 436 LRHPNCKQNL 448
 DB 716 LRNQSCKQCK 726

RESULT 13

PYA6_HUMAN

10 PYA6_HUMAN STANDARD; PRT: 1033 AA.

AC P59045; Q8NBF5;

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE PYRIN-containing APAF1-like protein 6.

GN PYPAF6.

OS Homo sapiens (Human).

CC Eukaryota; Metazoa; Vertebrata; Euteleostomi;

CC Mammalia; Eumetazoa; Primates; Catarrhini; Hominoidea; Homo.

CX NCBI:TaxID:9606;

RN (1)

SEQUENCE FROM N.A. (ISOFORM 1):
MEDLINE=22162427; PubMed=22019269;
Wang L., Manji G.A., Grenier J.V., Al-Gatwawi A., Wehrman S.,
Lora J.M., Geddes B.J., Briskin M., DiStefano P.S., Bertin J.,
"PyPA2", a novel p38 γ -containing Apaf-1 like protein that regulates
activation of NF-kappa B and caspase-1-dependent cytokine
processing.";
J. Biol. Chem. 277:29874-29882 (2002).
[2]
SEQUENCE FROM N.A. (ISOFORM 1):
TSSUS=247787;
MEDLINE=22388257; PubMed=12477793;
Strausberg R.L., Feingold E.A., Grouse L.H., Bergo M.O.,
Klausner R.D., Collins F.S., Wagner L., Shenoy D., Schuler G.D.,
Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
Hogkins R.F., Jordan H., Moore T., Vax S., Wang J., Hsieh F.,
Chacko K.L., Marziani K., Farmer A.A., Rubin G.W., Hong L.,
Stapleton M., Soares M.B., Bonaldo A.F., Casavant T.L., Scheetz T.E.,
Brownstein M.J., Ustin B., Tomiyaki S., Lathrop P., Prange C.,
Rata S.S., Loquellano N.A., Peters G.J., Aurora R.D., Mullan S.C.,
Bosak S.A., McEwan P.J., McKernan K., Vallee A.A., Garavito R.M.,
Ricciardi S., Worley K.C., Hale S., Garcia A.M., Day L.J., Bulvik S.W.,
Villalón D.K., Kuzoy D.W., Sodergren E., Lu X., Ellis P.A.,
Bailey A., Helton E., Kottmann M., Vadan A., Rodriguez S., Sanchez A.,
Whitting M., Madan A., Young A.C., Stevchenko Y., Burdard G.G.,
Bakerley R.W., Touchman J.W., Green E.D., Liskson M.C.,
Rodriguez A.C., Grimwood J., Schnitz E., Myers R.W.,
Buetefield Y.S.N., Kryzanski J.J., Skalska J., Smolius D.E.,
Schneitz A., Schein J.E., Jones S.J.W., Marra M.A.,
"Generation and initial analysis of more than 15,000 full length
human and mouse cDNA sequences.";
Proc. Natl. Acad. Sci. U.S.A. 99:16499-16504 (2002).
[3]
SEQUENCE OF 365-1033 FROM N.A. (ISOFORM 2):
TSSUS=Galal; tumor;
Kawakita B., Suiyara A., Takemoto M., Suiyara T., Irie R.,
C-Saki T., Sato H., Ota T., Kakamatsu A., Ishii S., Yamaguchi C.,
Isoro Y., Kawai-Hio Y., Saito K., Nishikawa T., Kimura K.,
Yamashita H., Matsuo K., Nakamura Y., Sekine Y., Kikuchi H., Kanda K.,
Wagatsuma M., Murakawa K., Kanohori K., Takahashi-Fujii A., Oshima A.,
Suzuki Y., Sugano S., Nagahara K., Kasahara Y., Nagai K., Isogai T.,
"NEO human cDNA sequencing project.";
Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
[4]
ALTERNATIVE PRODUCTS:
Event-Alternative splicing; Homo sapiens.
Name-1:
IsoId P59045-1; Sequence Displayed:
Name-2:
IsoId P59045-2; Sequence-NSP domain;
Note-No experimental confirmation available.
[5]
SIMILARITY: Contains 1 DAPIN domain.
[6]
SIMILARITY: Contains 6 leucine-rich repeat repeats
[7]
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL database
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed, usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
or send an email to license@sib-sib.ch).
[8]
EMBL; AY095145; AAM14632.1;
EMBL; BC034730; AAH34730.1;
EMBL; AK090621; BAC03490.1; ALT_INIT;
InterPro; IPR037091; LRR_RNinh;
InterPro; IPR033590; LRR_RNinh_sub;
InterPro; IPR037111; NACHT_NTFase;
InterPro; IPR034020; PAAD_DAPIN_dom;
Pfam; PF00560; LRR; 1;
Pfam; PF02758; PAAD_DAPIN; 1;
SMART; SM00369; LRR_R1; 9;
PROSITE; PS50824; DAPIN; 1;

PROSITE; PS50817; NACHT; 1;
ATP-binding; leucine-rich repeat; Repeat; Alternative splicing.
KW DOMAIN 1 91 DAPIN;
FT REPEAT 147 473 NACHT;
FT REPEAT 549 611 LRR 1;
FT REPEAT 632 655 LRR 2;
FT REPEAT 745 788 LRR 3;
FT REPEAT 802 827 LRR 4;
FT REPEAT 849 882 LRR 5;
FT REPEAT 919 944 LRR 6;
FT NP_BOND 153 158 ATP_INTERRUPTAL;
FT VARSPPLIC 614 688 SOMKSLVYKREKCSLYFTYMSRELHIFNDLNGISERILS
KALEHSKRLK -> R (in isoform 2);
FT CONFLICT 788 788 T -> I (in REF. 3);
FT CONFLICT 788 788 T -> I (in REF. 3);
SQ SEQUENCE 1031 AA; 117794 MW; A2580C485FBA49AC CRC64;
Query Match 18 73; Score 428.5; DB 1; Length 1033;
Best Local Similarity 26.23; Pred. No. 5.6e-28;
Matches 11; Conservative 85; Mismatches 211; Indels 71; Gaps 12;
QY 1 MKGATEKSMQQLIFSCWPEPSAPQELIRVERLDFIDGDELKPSFHDQPGWCLW 60
DB 19: INQ-KTSSSLAELIAKDWKGGAPADLSCKKLFLELDLN:RFEINNNESALCSN 249
QY 61 EKKPTTELLNSIRKKKGLPELSLITTPRTALEKHLLEHPRHVEILGISEARKEYF 120
DB 250 TORVPIPLVLVSLKKKMAPGWFTISRPTRGNKVKFLKEVDCTTLQLSNGKREIYF 309
QY 121 YVYHNAAAGQVFNVEENELPFGFVPLVGMVVCTCLOOLEGGGLRPTISITTTAV 180
DB 310 NSFFKQQRASAAQLVDEETILVGLCRVAILCWITCVLKSQDKGKGFQCCQTFDL 369
QY 181 YMYVLLSLGKPKGAFRLQFPNQ-----RGLCSLAADGLMNKILFEEQPLRKHGLDG 234
DB 370 HAFPLADATSEAG-----TANQVHLGLKXLLAGGLFSLTNSSGDLRCVGPT 424
QY 235 EVSFAELKNKIP-----QKD-----INCERSFLAT-----SR----- 262
DB 425 ADVSVLIQAANILLPSNTHRDYKFIHLNQVQFCTAIFLMAVNVYLIPSGSREYKREK 484
QY 263 -----FLEMLNEETRSHLEKSLCKWKFHKMDLW-----LQSKASDASTLQ 309
DB 495 YSDFRCVFTFEGLLNANRKKLETSPGYQCP---MVDSFKWYSGYMKKHLDPEKILTH 541
QY 210 KSLDEFFSYVECEEFICQALSHFOVIVYSNTASKMHWSSFCUKCRSAQVHLHYGA 369
DB 542 HMEFLFYCLYENREFFVKTVDALWEVTV-YLQSKOMVSLYCLDYCHLRTLKLSVQ 599
QY 370 TVSANUEDPARCSAGAHLLVQLPERTVLDAYSSEHLAALCTNENLIELSYRNALGR 429
DB 600 RLPQKKEP:LRPTASCKSLV-----YVREICSLFYTMESLRELHIFDNDLNGI 648
QY 432 GVKLLQCG:RHPNCKLQNL 449
DB 649 SERLSKALEHSSCKRLTK 668
RESULT 14
CARD_HUMAN STANDARD; PRT: 1040 AA.
AC Q9HC29; Q96RH5; Q96RH6; Q96RH9;
DT 28-FEB-2003 (Ref. 4); Created;
DT 28-FEB-2003 (Ref. 4); Last sequence update;
DT 15-SEP-2003 (Ref. 4); Last annotation update;
DE Caspase recruitment domain protein 15 (Nod2 protein); (Inflammatory
bowel disease protein 1);
GN CARD15 OR NOD2 OR IBD1;
OS Homo sapiens (Huma);
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo;
OX NCBI_TaxID:9606;
RN [1]

SEQUENCE FROM N.A. (ISOCFORMS 1 AND 2). METAGENESIS OF LYS 305, AND
VARIANT ARG-938.
TISSE:Breast;
RX PubMed:11587742;
RA Gupta V., Itohara N., Benito A., Chen P.H., Yutaka S., Nunez G.,
"Nod2, a Nod1/Api-1 family member that is restricted to monocytes and
activates NF-kappaB",
Acta Biol. Chem. 276:4812-4818(2001).
RE
RE
SEQUENCE FROM N.A. (ISOCFORMS 1 AND 2). VAPANTS CD THR-140, ARG-157;
CYS-235, ARG-248; SEP-269; SEP-347; ASN-201; VAL-311; TRP-311;
VAL-348; ARG-352; CYS-373; SEP-414; LEU-441; VAL-441; LYS-441;
VAL-622; TRP-612; TRP-654; TRP-724; CYS-747; CYS-757; GLY-725;
VAL-755; VAL-758; LYS-778; MET-793; CYS-843; SEP-843; VAL-863;
ARG-908; ASP-918; ASP-924 AND ILE-955, VARIANTS ULCERATIVE COLITIS
THR-140 AND THR-865, AND VARIANT MET-159.
TISSE=Leukocyte;
MEDLINE=21279172; PubMed=11355576;
RA Bugot G.-P., Chamillard M., Zouali H., Lesage S., Gerard C.-P.,
Belache G., Almer S., Tysk C., McNair C.A., Gassart M., Binder V.,
Finkel Y., Cottet A., Modigliani R., Laurent P.J.P.,
Gower-Rousseau C., Macry C., Colombel J.F., Saharou M., Thomas G.,
"Association of NOD2 leucine-rich repeat variants with susceptibility
to Crohn's disease",
Nature 411:599-603(2001).
RE
RE
[3]
VARIANTS BS GLN-334; TRP-334 AND ILE-453.
VAL-114;944; PubMed=1158764;
RA Michel P., Richard C., Lesage S., Rykaczak M., Hleut A.,
Mandouvier-Hard S., Hafner R., Chamillard M., Zouali H., Thomas G.,
Bugot G.-P.,
"CARD15 mutations in Blau syndrome",
Nat. Genet. 20:19-20(2001).
RE
RE
- FUNCTION: Induces NF-kappaB via RICK (CARD14, RIP1) and IRE
gamma. Confers responsiveness to intracellular bacterial
lipopolysaccharides (LPS).
- SUBUNIT: Binds to RICK by CARD CARD interaction.
- SUBCELLULAR LOCATION: Cytoplasm.
- ALTERNATIVE PRODUCTS:
Event: Alternative initiation.
Comment=2 isoforms. (Nod2) (known beta-1 and CARD20, are
produced by alternative initiation with isoforms can activate
NF-kappaB. Isoform 1 is more specific.
- TISSUE SPECIFICITY: Monocytes specific.
- DISEASE: Defects in CARD15 are the cause of Blau syndrome (BS),
(MIM:166540) a rare autosomal-inherited febrile childhood-onset
early-onset granulomatous inflammation, arthritis, uveitis and skin rash.
- DISEASE: Defects in CARD15 are the cause of susceptibility to
Crohn's disease (CD) (MIM:160910), a chronic inflammatory bowel
inflammatory bowel disease. It may involve any part of the
gastrointestinal tract, but is most frequently the terminal ileum and
colon. Bowel inflammation is transmural and discontinuous. Crohn's
disease is commonly classified as inflammatory disease.
- DISEASE: Defects in CARD15 are the cause of susceptibility to
ulcerative colitis (MIM:160910), a chronic inflammatory bowel
disease. In ulcerative colitis, the inflammation is continuous and
limited to rectal and colonic mucosal layers. Ulcerative colitis
is commonly classified as inflammatory disease.
- SIMILARITY: Contains 2 CARD domain.
- SIMILARITY: Contains 1 NACHT domain.
- SIMILARITY: Contains 6 leucine-rich LRR repeats.
This SWISS-PROT entry is copyrighted. It is included through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as the copyright notice is in no way
modified and this statement is not removed. Please refer to the Commercial
entities requires a license agreement. See <http://www.ebi.ac.uk/submit/embout/>
or send an email to license@ebi.ac.uk.

GenCore version 5.1.6
Copyright (c) 1993 - 2001 GenPhen Ltd.

CV Protein - protein search, using sw model

Run on: October 31, 2003, 06:54:44, Search time 114 Seconds
(without alignment)
1016.365 Million cell updates/sec

Title: US-10-028 374 2

Perfect score: 2377

Sequence: 1 MNQATECSQQLIFSCWPE.....GVKLVGGLSHPRKCNLR 449

Scoring table:

BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 830525 seqs, 256052604 residues

Total number of hits satisfying chosen parameters: 840525

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

```

1: sp_archaea;
2: sp_bacteria;
3: sp_fungi;
4: sp_human;
5: sp_invertebrate;
6: sp_mammal;
7: sp_rhco;
8: sp_organellae;
9: sp_phage;
10: sp_plant;
11: sp_rodent;
12: sp_virus;
13: sp_vertebrate;
14: sp_unclassified;
15: sp_virus;
16: sp_bacteriap;
17: sp_archaeap;

```

Fred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	DB ID	Description
1	2334	98.2	892	4	Q8NEU4	Q8NEU4 homo sapien
2	1145	48.2	1035	11	Q8R4B6	Q8R4B6 mus musculus
3	594	25.0	713	6	Q95127	Q95127 macaca fasc
4	563	23.7	982	11	Q8R440	Q8R440 mus musculus
5	549.5	23.2	846	4	Q8IX10	Q8IX10 homo sapien
6	519	21.8	963	11	Q8CCJ9	Q8CCJ9 mus musculus
7	494	20.8	748	11	Q9NMW0	Q9NMW0 mus musculus
8	472.5	19.9	673	11	Q8CCK1	Q8CCK1 mus musculus
9	443	18.6	825	11	Q8C6M5	Q8C6M5 mus musculus
10	393.5	16.6	657	11	Q9EP37	Q9EP37 mus musculus
11	261	11.0	953	4	Q8WFS5	Q8WFS5 homo sapien
12	253	10.6	1020	11	Q8K120	Q8K120 mus musculus
13	246.5	10.4	778	4	Q8NF48	Q8NF48 homo sapien
14	233	9.8	726	11	Q8BTJ6	Q8BTJ6 mus musculus
15	233	9.8	953	11	Q8B580	Q8B580 mus musculus
16	223	9.4	660	4	Q8NF66	Q8NF66 homo sapien

17	212.5	8.9	519	11	Q8C243	Q8C249 mus musculus
18	212.5	8.9	864	7	Q29675	Q29675 homo sapien
19	212.5	8.9	932	4	Q96K14	Q96K14 homo sapien
20	200	8.4	682	4	Q96D51	Q96D51 homo sapien
21	176	7.4	1052	7	Q9GCD8	Q9GCD8 rattus norv
22	176	7.4	1033	7	Q9GCD9	Q9GCD9 rattus norv
23	176	7.4	1153	7	Q9GDE0	Q9GDE0 rattus norv
24	174	7.3	977	7	Q8HW89	Q8HW89 mus musculus
25	174	7.3	1105	7	Q9TFP1	Q9TFP1 mus musculus
26	165.5	7.0	516	4	Q8BFF5	Q8BFF5 homo sapien
27	165.5	7.0	1056	4	Q8NF42	Q8NF42 homo sapien
28	155	6.5	743	4	Q8TEE2	Q8TEE2 homo sapien
29	150.5	6.4	1064	5	Q94429	Q94427 ciona intes
30	118	5.8	1087	4	Q9H6Y0	Q9H6Y0 homo sapien
31	135.5	5.7	287	4	Q9H125	Q9H126 homo sapien
32	135	5.7	504	4	Q9H528	Q9H528 homo sapien
33	116.5	4.9	1075	5	Q94430	Q94430 ciona intes
34	113.5	4.8	392	11	Q8BWC2	Q8BWC2 mus musculus
35	111	4.7	269	11	Q9D458	Q9D458 mus musculus
36	110	4.6	456	11	Q91V17	Q91V17 mus musculus
37	109	4.6	393	5	Q81FT8	Q81FT8 marthasteri
38	108	4.5	461	6	Q8H2P4	Q8H2P9 pan troglod
39	107	4.4	447	4	Q96FD7	Q96FD7 homo sapien
40	107	4.4	461	4	Q9HC80	Q9HC90 homo sapien
41	107	4.4	461	4	Q84ZK8	Q84ZK8 homo sapien
42	107	4.3	2550	12	Q8BEN2	Q8BEN2 cecropouche v
43	106.5	4.3	390	13	Q9HT12	Q9HT12 xenopus lae
44	106.5	4.3	1160	4	Q84Z24	Q84Z24 homo sapien
45	106	4.3	456	11	Q924P4	Q924P4 mus musculus

ALIGNMENTS

RESULT 1

Q8NEU4 Q8NEU4 PRELIMINARY; EPT; 892 AA.
AC Q8NEU4;
CT 01-OCT-2002 (TREMURel. 22, Created:
CT 01-OCT-2002 (TREMURel. 22, Last sequence update:
CT 01-MAR-2003 (TREMURel. 23, Last annotation update)
DE Monarch:1 splice form IV.
CS Homo sapiens (Human)
CC Exarystora; Metazoa; Chordata; Craniata; Vertebrata; Eutelesstomi;
CC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
CX NCBI: TaxID:9606;
RN 11;
RP SEQUENCE FROM N.A.
RA Williams K.L., Lischoff M.W., Tang J.P.Y.;
RI Submitted (MAY 2002) to the EMBL/GenBank/DBJ databases.
DR ENBL; A116207; AAK75145.11;
DR InterPro; IPR007091; LRR_RNrich.
DR InterPro; IPR007111; NACHT_NTPase.
DR InterPro; IPR004020; PAAD_DAPIN_dom.
DR Pfam; PF02758; PAAD_DAPIN_1.
DR PROSITE; P550824; DAPIN; 1.
DR PROSITE; P550503; LRR_R1; 1.
DR PROSITE; P550817; NACHT; 1.
SQ SEQUENCE 892 AA; 101733 MW; AE703D8DF341C2AC CRC64;

Query Match: 98.2%; Score 2334; DB 4; Length 892;
Best Local Similarity 90.7%; Pred. No. 2e-218;
Matches 449; Conservative 0; Mismatches 0; Indels 46; Gaps 2;

QY	1	MNQATECSQQLIFSCWPE	SAPQLIRVPERLLFI	DGDELKPSFHD	QGGWCLQW 60
DB	255	MNQATECSQQLIFSCWPE	SAPQLIRVPERLLFI	DGDELKPSFHD	QGGWCLQW 314
QY	61	EKKRTTEKNSLRKKKLP	ELSLCTTRPTALEK	HLRLLHPRHVEL	IGFSEARKEYF 120
DB	315	EKKRPTTEKNSLRKKKLP	ELSLCTTRPTALEK	HLRLLHPRHVEL	IGFSEARKEYF 374
QY	121	YKYPHNASAGQVFNVR	NEPLFMCVFPLVCW	VVCTCLOQQJEGG	LLRTSRTTAV 180

```

DB 375 YKFEHNAEQGVNRYVRNREPLDTWTFVPLVQVWVWVTCQCCQLEGGLRQTSRTTAV 434
181 YVYLLSNOCKPKAPRLQPPFNQPGGLCSAAADGKWKTKLIFERQTKRKGGLDGEDVSAP 243
415 YVYLLSNOCKPKAPRLQPPFNQPGGLCSAAADGKWKTKLIFERQTKRKGGLDGEDVSAP 494
241 LNMKIFQKIN 255
495 LNMKIFQKIN 554
256 FLALTSRFLPGLLNEERFSHLEKSLQKVSQPHLEKMLQNLQNSKQAGDGTQQGSLLEFF 315
555 FLALTSRFLPGLLNEERFSHLEKSLQKVSQPHLEKMLQNLQNSKQAGDGTQQGSLLEFF 614
316 SOLVEIOEEEFIQALSHFQVIVVSNIAASKMHEVVSFCIKPERSAQVHLYGATYSADG 375
615 SOLVEIOEEEFIQALSHFQVIVVSNIAASKMHEVVSFCIKPERSAQVHLYGATYSADG 674
376 EDRARCAGAGHTLAVQQLPRETVLIDAYSEHLAALQNLNFIHLISYRNALSGRGYKOL 434
675 EDRARCAGAGHTLAVQQLPRETVLIDAYSEHLAALQNLNFIHLISYRNALSGRGYKOL 734
415 CQGLRHKNCKLQNL 449
735 CQGLRHKNCKLQNL 749

RESULT 2
Q95LZ7 PRELIMINARY; FRT; 1033 AA.
AC Q95LZ7
DT 01-JUN-2002 (TRENHrel; 21, Created)
DT 01-JUN-2002 (TRENHrel; 21, Last sequence update)
DT 01-MAR-2003 (TRENHrel; 23, Last annotation update)
DE Mast cell maturation inducible protein 1.
OS Mus musculus (Mouse).
CC STRAIN:BAE/CJ;
CC Kikuchi-Yanosita R., Koga K., Taketomi Y., Taniguchi I., Saito T.,
CC Ishii S., Hisada Y., Suzuki N., Shimizu T., Hanaoka M., Kato T.,
CC Chao H.-W., Sawada M., Inagaki N., Nagai H., Yoshida M., Kudo I.,
CC "Identification of inducible genes during in vitro maturation of mouse
CC bone marrow-derived mast cells to connective tissue type mast cells."
CC Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF486632; AAL90974.1;
DR InterPro; IPR001592; HLM Basic
DR InterPro; IPR001511; LRR
DR InterPro; IPR007391; LRR RNase
DR InterPro; IPR003590; LRR_PNICH_sub
DR InterPro; IPR007111; NACHT_NTPase
DR InterPro; IPR004320; PAAD_CAPIN_3m
DR Pfam; PF02560; LRR; 2
DR Pfam; PF02758; PAAD_CAPIN; 1
DR SMART; SM00369; LRR_E; 1
DR PROSITE; PS50824; CAPIN; 1
DR PROSITE; PS00039; HLM 1; 1
DR PROSITE; PS50503; LRR R1; 3
DR PROSITE; PS50837; NACHT; 1
SQ SEQUENCE 1033 AA; 118274 MW; 5924690366b11177 PCF4;

```

Query Match 49.2%; Score 1145; DB 1; Length 1033;
 Best Local Similarity 45.2%; Pred. No. 2,2e-102;
 Matches 231; Conservative 83; Mismatches 127; Indels 70; Gaps 5;

QY 6 TCSNQCILFSCWPSAPLQGLVWPRRLFLIISGFLKPSHDPQPTNCLCWEKRRP 65
 DB 264 TPRSACLIIVSWPDPNPVWIKKKGRLLFMDGFTLQAFLEHICVCTNCKAVR 323

```

QY 66 TELLNSARKKZKZELFUSLITTRTALAEK:HRLLERPHRVILGFSBAERKEYFYKYFH 125
DB 324 GQILLSSLRKKLLDPKASLITTRVALEKQLHLLDHPRHVEILGFSBAERKEYFYKPS 383
QY 126 NAEQACVENVYRDNHPLFTKFEVGVWVWVTCQCCQLEGGLRQTSRTTAVMYLV 165
DB 384 NELQAREAPLQREVEVFTMGFPLVVMVITGLQQMETGKSJQTSKTTTAVYVFFL 443
QY 186 LSNQPKGAPRLQPPFNQPGGLCSAAADGKWKTKLIFERQTKRKGGLDGEDVSAPLNMKI 245
DB 444 SSLQSPGGLSEHLEKSLQKVSQPHLEKMLQNLQNSKQAGDGTQQGSLLEFF 503
QY 246 FQKQINER 254
504 FQKQINER 563
255 -----SFLALTSRFLPGLLNEERFSHLEKSLQKVSQPHLEKMLQNLQNSKQAGDGTQQ 309
564 GKPKYKILFVVRFLGLVYNQERTSYLKKLSCKISQVRDELLKMWIEVKAKKLQWOP 623
310 GSLEPFTSIVECEERTIQCALSHFQVIVVSNIAASKMHEVVSFCIKPERSAQVHLYG 369
624 SOLEFVCLVEMQEEVAVQSMIHFKILNLSRYDHWVVSFCIKNCHRYKLTSLGF 581
370 TYSAGELNARCASAGHTLAVQQLPRETV-----LIDAYSEHLAALQNLNFI 417
682 FNSPKXEBEERPGGRPLDQVQVFPDTHVASSRLVNCCLTSSFCRGFSSLSLTPSIT 741
418 ELSLYNALSGRGYKOLQGLRHKNCKLQNL 448
742 ELSLYNALSGRGYKOLQGLRHKNCKLQNL 772

RESULT 3
Q95LZ7 PRELIMINARY; FRT; 713 AA.
AC Q95LZ7
DT 01-DEC-2001 (TRENHrel; 19, Created)
DT 01-DEC-2001 (TRENHrel; 19, Last sequence update)
DT 01-MAR-2003 (TRENHrel; 23, Last annotation update)
DE Hypothetical: 80.3 kDa protein.
OS Macaca fascicularis (Orang-eating macaque) (Cynomolgus monkey).
CC Exaiyoka, Metazoa; Chordata; Mammalia; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
CC Cercopithecoidea; Macaca.
CC NCBI TaxID=9544.
RN [1]
RP SEQUENCE FROM N.A.
RC TSSUE=Testis;
RA Hashimoto K., Osada N., Hiji M., Kusuda J., Tanuma K., Hirai M.,
RA Terao K., Sugano S.;
RT "Isolation of novel full-length cDNA clones from macaque testis cDNA
RT libraries."
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB011044; BAB66437.1;
DR InterPro; IPR007097; LRR RNase
DR InterPro; IPR007111; NACHT_NTPase
DR PROSITE; PS50503; LRR R1; 2
DR PROSITE; PS50837; NACHT; 1
KW Hypothesized; Protein;
SQ SEQUENCE 713 AA; 80259 MW; 6F214C9B773F54DC CRC64;

```

Query Match 25.5%; Score 594; DB 6; Length 713;
 Best Local Similarity 28.5%; Pred. No. 7e-49;
 Matches 162; Conservative 81; Mismatches 174; Indels 154; Gaps 13;

QY 13 LYSWPEISAPQLRIPERLLFIIDFDELKPSFHDPOGW-----CLCWEKRRPTL 68
 DB 1 XVFEDPEKCDIFHLAQACKLTFVIGDFDELGA----PPGALIQICGWEQKQKPVV 56
 QY 69 LNSLSEKRLIPESLITTRPTALEKLRLLDHPRHVEILGFSBAERKEYFYKYFHNAE 128

DB	53	UAGSILKRRKMLPKATLLVTRPRAKRIIPFLAEQIVIRVLEHEDRRAVYELPHIGDEE	116
QY	129	QACQVENVYRDNEFIETMGVFLVWVSTTQQLUESGILPQDPTTAVYKVLASL	188
DB	127	QAMRAELMRSMNAALFQLGSAPAVWVYVTTKQDMERHEDAPTITLSTRLFPDLS	175
QY	189	MQPKGAPRQPPNQRGKLSIANGLKILFEGVPRKRIKIDGVSANFKNIPCK	248
DB	176	QPPQA---QLGALRAISLNAUSGKACMSVHHELEKAGVLESDRIPLDRIIDRC	231
QY	249	D-----INDESFAL-----	259
DB	232	QGVAKGCVYFNLSPQFLTALFYALEREEDRDHAKLIIVKGLASPEELKPKDL	291
QY	260	TSRPLGGLINBETHRLERSLQWVSHIKMLLQWISKKAQSDSTLQQSLFFSC	318
DB	292	QACRFIFGLANKRVKLEATEKWNSEFEIKQLIRCVMSFKHHTNADIREILCCL	349
QY	319	YEIOEEFICQALSHFCVIVWSNIASKYEMWNSFYLAKRS-----A	361
DB	350	YESQDELYKVMQPKELSHIMAYDLSFQFRCSCUIFFADACRNICLALRGHK	407
QY	362	QVHLAYGATYSADGDR-----A-CSA-----	383
DB	408	TVTHL-----TLOGTQKXNGALGVHPRPNLLEYGLVADSATQCAWLSIALKAN	462
QY	484	-----GATILVQCP-----ENTVLDAYSEHLAAALTNKRL	416
DB	463	RSIMVNLSDNELDEGAKLTLTHPKCEQFLDENHETFAKCKGLAAVLYSPRL	522
QY	417	ILSLVYNALGSPGVKLAQVIRHNRKIQNL	448
DB	523	THLTIANKSKRTGVAFKFGHLYETFFCAL	564
RESULT 4			
Q8404			
ID	Q8404	PRELIMINARY:	187; 962 AA.
AC	Q8404		
DT	01-MAR-2003	(TEYBRLrel, 23, Created)	
DT	01-MAR-2003	(TEYBRLrel, 23, Last sequence update)	
DT	01-MAR-2003	(TEYBRLrel, 23, Last annotation update)	
DE		Similar to PAN2 protein.	
CS		Mus musculus (mouse).	
CC		Eukaryota; Metazoa; Chordata; Mammalia; Artiodactyla; Eulipotyphla.	
CC		Mammalia; Eutheria; Rodentia; Muridae; Murinae; Murinae; Murinae.	
CC		NCBI TaxID:101090.	
RN		[1]	
RP		SEQUENCE FROM N.A.	
RC		STRAIN=CS7BL/6J; TISSUE=Ovary;	
RX		NEBI:NEB2354693; PubMed:12466411	
RA		The RIKEN Consortium.	
PA		"Analysis of the mouse transcriptome based on high-speed annotation of the RIKEN Genome Expression Research Group Phase 1 & II Team."	
RC		62,720 total length cDNAs."	
RI		Karlsruhe 4231561-573126023.	
TR		EMBL: AK07843; BAC400241; ...	
SQ		SEQUENCE 982 AA; 112600 MW; 682RCPEALQVLSRS P 64.	
Query March 23.74; Score 563; DP 11; Length 982;			
Residual Similarity 27.0%; Pred. No. 116 45;			
Matches 143; Conservative 136; Mismatches 147; Indels 94; Gaps 9.			
QY	9	SVQSLQSWPEPSAPQLQIRVFERLCITIDIDRLKLSFQCPQPKLQWEEKRPTL	68
DB	193	SAELISRWKPGDSAEILSKPEKALITLSDKNEELIYKSEKCONTEKQPVAV	258
QY	69	LKSLIRKLLPELSLITITRTALKSKRLHSHPRNVEIDTESPAERKRYFYKPHNA	127
DB	259	LSSLLRKLKLPESLSLISATFESKVENRIEYTHGRIKGLAKRNKMSPHRLFQIR	317
QY	128	EQACQVENVYRDNEFTMGVFLVWVSTTQQLUESGILPQDPTTAVYKVLASL	187

DB	318	NEAHFENFNRNENQIFTVQVPRKQKMWATC:KEELKGRDPVSIQKCTTSLSLTHFEN	377
QY	188	LMQKRPDAPLQPFPRQKGLCSAAAGLWKNQKILFEDQDLKRGDGEVSFAFNXNIF	246
DB	178	LEITQNAHSEKSKQDQLOHLSAAAGMWDTFVFGEEALRRNSIMDSQIRILLQVYL	437
QY	247	QKDNCEHSEFAL-----	260
DB	416	INDEHNSYTFHNSVEVCAAFYLLKSNIVHSPSEVKSIEKIMFAFKKVVQKIFF	497
QY	261	SREPLNENHSTRSHKSLQWVSHIKMLLQWISKKAQSDSTLQQSLFFSC	320
DB	418	OSFPLGLIHSEQKLSAPRHQLSCEIKRQVQCLSTISQNESLQEGIDMKILFYCLPE	557
QY	331	IQEELFQALSHFCVIVWSNIASK--MERVSSFCVKRCESQVHLHYGATYSADGDR	378
DB	558	MDDDTFVANKKVEQI---NFAKVCYSDVVAARCLKHCFTLKLKLSFSTQKVLGQAEH	614
QY	379	A---RCSAGATILY-----QLPRTVLIDAYS-----	403
DB	615	SYVEFLDAMKHEIGSVFTISKIQELRMKQTNLSASFVLYNKLKYHNYTINVLVANNV	674
QY	404	---HELAALCNPMLHLSGVNALGSRGVKILCOGLRHHPCKIQNL	445
DB	675	PRVLEKVFPRGLKWNKCHLNLSITLSISQVLLCDVLSQAFQNTIED	724
RESULT 5			
Q81XTO			
ID	Q81XTO	PRELIMINARY:	187; 846 AA.
AC	Q81XTO		
DT	01-MAR-2003	(TEYBRLrel, 23, Created)	
DT	01-MAR-2003	(TEYBRLrel, 23, Last sequence update)	
DT	01-MAR-2003	(TEYBRLrel, 23, Last annotation update)	
DE		Similar to NALP2 protein.	
CS		Homo sapiens (human).	
CC		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
CC		Mammalia; Eutheria; Primates; Catartida; Homnidae; Homo.	
CC		NCBI TaxID:9606.	
RN		[1]	
RP		SEQUENCE FROM N.A.	
RC		TISSUE=Testis.	
RA		Strasbourg F1.	
RI		Submitted (May 2002) to the EMBL/GenBank/CCDS databases.	
TR		EMBL: BC040000; GenBank: AF042641.	
SQ		SEQUENCE 846 AA; 10164 MW; 4788CLASGQSLQPT CAS74.	
Query March 23.74; Score 549.5; DP 4; Length 846.			
Residual Similarity 30.1%; Pred. No. 196 44;			
Matches 151; Conservative 85; Mismatches 181; Indels 86; Gaps 13.			
QY	8	CSVQSLQSWPEPSAPQLQIRVFERLCITIDGFDLKRSHDQGPQPKLQWEEKRPT	67
DB	256	CSFADENRWKPELQCDRIPLKAKKILFVLDGFDLGAAPGALIEDICQENKRRKVP	215
QY	68	GLNSLIRKLLPELSLITITRTALKSKRLHSHPRNVEIDTESPAERKRYFYKPHNA	127
DB	316	VLLQSLNRRVLPKALIVITRPEADPLCLLAEERIVIRVEGFLDEEDRAYFLRHGDE	375
QY	128	EQACQVENVYRDNEFTMGVFLVWVSTTQQLUESGILPQDPTTAVYKVLASL	187
DB	376	EQAEAFELMPSNAALFQLGSAPAVWVYVTTKQDMERHEDAPTITLSTRLFPDLS	425
QY	158	LMQKRPDAPLQPFPRQKGLCSAAAGLWKNQKILFEDQDLKRGDGEVSFAFNXNIF	247
DB	416	PF---EQACQVENVYRDNEFTMGVFLVWVSTTQQLUESGILPQDPTTAVYKVLASL	490
QY	248	KD--INDEHNSYTFHNSVEVCAAFYLLKSNIVHSPSEVKSIEKIMFAFKKVVQKIFF	264
DB	411	QCRVAKGTYSFTHSGLQCFQTAIFYTLKKEEELRQHTMCIQSDGVKLLSGVERLNPD	549
QY	255	EQACQVENVYRDNEFTMGVFLVWVSTTQQLUESGILPQDPTTAVYKVLASL	314

```

DB 550 LQAGYVSYGLANKEKRAKELEATEVEMSPDIFQKELLREIDIKROGHSVTVDQ ..... 605
QY 315 FSLVIEIQEERTQOALSHPQVWVSNIAKNEHYVSSSELZECUSAQVHLV ..... 167
DB 626 LGLYSEOEELVKEVYAFKKE: SUBNANVPSSTVYKRNILQKYSQVAKENLP 663
QY 468 GATYSADHEDPARCSAGANTLAVCLPARKVILDAYSHHAAALCTNPKLLELSVSNAL 426
DB 664 ENVTASEDAEVSQDQCH: KLPPFTGLCSFG ..... 711
QY 427 GSRGVKGLAOGLRHPNKKCNL 448
DB 712 SASLVLCEQIASOTCHLGRV 733

RESULT 6
Q99XWC PRELIMINARY; PST: 561 AA.
AC Q99XWC:
DT 01-MAR-2003 (Tremblere). 23, Created:
DT 01-MAR-2003 (Tremblere). 23, Last sequence update:
DT 01-MAR-2003 (Tremblere). 23, Last annotation update:
DE Hypothetical RN1-like structure containing protein.
OS Mus musculus (Mouse).
CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI TaxID:10090;
RN 1.
RP SEQUENCE FROM N.A.
PC STRAIN=C57BL/6J; TISSUE=OVARY;
RX MEDLINE=22354681; PubMed=12456481.
RA The FANTOM Consortium.
EA the RIKEN Genome Exploration Research Group Phase 1 & 2 Team.
RT "Analysis of the mouse transcriptome based on functional annotation of
PT 62,770 full-length cDNAs".
RL Nature 420:561-573 (2002).
DR EXBL: AK054426; SACS5775.1.
KW Hypothetical protein.
SQ SEQUENCE 561 AA; 99955 MW; D1A6A79A2A44F E 6347

Query Match 21.48; Score 519; DB 11; Length 561;
Best local similarity 24.68; PId: No. 4.2e-39;
Matches 149; Conservative 91; Mismatches 160; Indels 180; Gaps 10;

QY 9 SSKLIFSCWFEESAPQCELLIAVRELPITDIPGFIPTFHEDVGFEMFLCWEEKPFTS 68
DB 194 SLAMTSRWPSPSAPMEETISCHMILPTDLEAVKATLQGLQVWKEKGVNV 253
QY 69 LLSLIRKKLLPELSLITRTTALKEHLEHLEHLEHLEHLEHLEHLEHLEHLEHLEH 128
DB 254 LLSLIRKKLLPELSLITRTTALKEHLEHLEHLEHLEHLEHLEHLEHLEHLEHLEH 313
QY 129 GAGGVNVRDNPGLFTMGVPLVGVWVWVWVWVWVWVWVWVWVWVWVWVWVWVWVW 188
DB 314 LAGAFSLRNECLFTVCGAPVWVWVWVWVWVWVWVWVWVWVWVWVWVWVWVWVW 373
QY 189 MQKPGAPRLQPPNCR: GLASAADEKSKKIDLELDELPLKPLGIDGVSAFLRM 243
DB 374 FIFR: NAQPSNNSDLENLQFVAVEMATVAVENALRANGVQSDIFITLDDI 429
QY 244 NIFQKQINCERSFAL: ..... 259
DB 430 GLEQSESENSY:FLRPSVQVRVNAAMYILHSETHSQVWVYFETPLFLNKKW 489
QY 260 -TSSEFLGLNEETRSHLEKSLKWKVSHIKMDLLQWISQKASQSDSTLQQ 309
DB 490 VFLOCFEFLGHETEQEKLEAFPHHLEKLEKLEKLEKLEKLEKLEKLEKLEKLEKLE 316
QY 317 CQYRQEEETIQQA: LSHFVIVVSNIAK: ..... 345
DB 550 CCFEMSEEVQSAVNCREQIDVYKGVVQVDFVNAVAVTSHGALITFSTSAQVLAKEIG 605
QY 346 ..... MEHVS 351

```

```

DB 610 QPKKILHWHQICSVF:PNKQIKTLRIEDTTFNEPVFKIFYSY:KNSC:IKTLVAVNS 669
QY 352 SFCLK: .....CRSAQVHLVYATVS ADGE: ..... LKRAKSAQHT 387
DB 610 PCKKSLKLEHILQYHNFELVLRTEFLSHSVFM:CTILNQAE:NIALLILANKSLCHS 729
QY 388 LVA21PRTVLLDAYSHHAAALCTNPKLLELSVSNALSGRVKLLCQSLRHPNKKLN 447
DB 930 W: ..... 774
QY 448 ..... 448
DB 775 ..... 775

RESULT 7
Q99XWC PRELIMINARY; PST: 748 AA.
AC Q99XWC:
DT 01-JUN-2001 (Tremblere). 19, Created:
DT 01-JUN-2001 (Tremblere). 19, Last sequence update:
DT 01-MAR-2003 (Tremblere). 23, Last annotation update:
DE Ribonuclease (angiotensin inhibitor) 2.
OS Mus musculus (Mouse).
CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI TaxID:10090;
RN 1.
RP SEQUENCE FROM N.A.
PC TISSUE=Testis;
RX MEDLINE=21757048; PubMed=12735251;
RA Wang P.J., McFarley C.R., Yang F., Page D.C.
RT "An abundance of X-linked genes expressed in spermatogenesis".
RL Nat. Genet. 23:422-426 (2001).
DR EXBL: AF235581; AKR1961.1.
DR HSSP: P10755; 2BRH.
DR MGI:1896018; Rn2.
DR InterPro: IPR007091; LRR_PNIN.
DR PROSITE: PS50503; LRR_R1.
SQ SEQUENCE 748 AA; 84906 MW; 674A445E01D58E8 CRC64;

Query Match 20.84; Score 494; DB 11; Length 748;
Best local similarity 26.33; PId: No. 4.2e-39;
Matches 129; Conservative 90; Mismatches 171; Indels 94; Gaps 9;

QY 52 LFWKEERETHEILNLSLIRKKLLPELSLITRTTALKEHLEHLEHLEHLEHLEHLEH 116
DB 133 QVETTERVPLILANSLRRNMLKSGSP:SATPTTFKMEGRVETCNVKTIVTFENSI 72
QY 117 KEVFKYFENRACQGVNVEVNEFLPTNGFVPLVGVWVWVWVWVWVWVWVWVWVW 176
DB 93 KMYPSL:QKRTKTKQFHESANKKQOFLTVGVPLVGVWVWVWVWVWVWVWVWVWVW 132
QY 177 TTAIVYVLSLMDPKKA: FLQPPVNCRCGLSLAUGLWVWVWVWVWVWVWVWVWVW 235
DB 133 TTSATTHIRPLFPQSAQYPSKESQALQSCSLAASQVMTDTFVFGFALRRNG:MDS 192
QY 236 DVSAFLNKNIFQKQINCERSFAL: ..... 259
DB 193 DPTPLDVLHLEKSKSKSKSYFLHPSITQEVCA:FYLLKSHKHPSQCVKSIALITFT 252
QY 260 .....SRPFGILNEETRSHLEKSLKWKVSHIKMDLLQWISQKASQSDSTLQQ 309
DB 253 LKKVAVGVNIFEGSPF:FLHLESEOKLEAFPHHLEKLEKLEKLEKLEKLEKLEK 312
QY 310 LSHFVIVVSNIAK: MEHVSFCLKPR:SAQVLAKEIG 367
DB 317 CQYRQEEETIQQA: LSHFVIVVSNIAK: ..... 345
QY 550 CCFEMSEEVQSAVNCREQIDVYKGVVQVDFVNAVAVTSHGALITFSTSAQVLAKEIG 605
DB 346 ..... 346

```

```

DB 370 TQNLSEGGHSEYTEKLMCNHMGSVLSSKD:Y:LOVKR:NIATASLVYS-HLXYP 428
QY 410 DCT .. .....NPNDLS:YRNAL:SGRYKJ:COG:PHNPKLQ 446
DB 420 SUTKALVWNVTFICNDRFTEL:QUCLO:LDLMLTFSHCWKLLCOWLSQEECNIE 488
QY 447 KQ 448
DB 489 KQ 490

RESULT 8
Q8CGN1 PRELIMINARY; PRT: 673 AA.
AC Q8CGN1;
DT 01-MAR-2003 (TREMblrel: 23, Created)
DT 01-MAR-2003 (TREMblrel: 23, Last sequence update)
DT 01-MAR-2003 (TREMblrel: 23, Last annotation update)
DE 2RAD and NACHT containing protein.
CS Mus musculus (Mouse).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
CX NCBI_TaxID:10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Olfactory brain;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium.
RA the RIKEN Genome Exploration Research Group Phase 1 & 11 Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RJ Nature 420:563-573(2002);
DR EMBL; AK032446; BAC27872.1;
SQ SEQUENCE 673 AA; 76367 MW; EE773C592BEC7054 CP664;

Query March 19, 98; Score 472.5; DB 11; Length 673;
Best Local Similarity 30.7%; Pred. No. 4.5e-37;
Matches 120; Conservative 75; Mismatches 119; Indels 77; Gaps 10;

QY 7 EGSMDQLFSCMPSPAPLOELRPVPERLFIIDGDELKSPHPQGWCLWEEKRPT 66
DB 212 KCSLPNLCWCCGDDQAFVTELR_FUCLDFILGVDRLQKSSR .....A 256
QY 67 ELLINS:LRKKLPELSLITTRPTALKRHLRHLHPRHPRHVEILGFSHAERK 126
DB 257 ECVHLIMRRREV:KSLITITFIALQS:EKML:EFRRVLMVFSEEDREYFSSCTD 315
QY 127 AKGAGQVNVVRNRP:FTWCVPVLYWV:TVLQGLCEHEDLQTSRTTAVVYLYL 156
DB 316 NEQLKNALEFQNNAVYKACQVPHQVWVWVWVWVWVWVWVWVWVWVWVWVWV 374
QY 187 SLVQPKRG:--APRLQFPNORGLS:AMGLWQKILFEQGLRKHGLDGEDVSAFLNM 243
DB 375 SIFPTDGGDSSELTRHKV:KSLCSLAAGRMARQRLFEEDVEVRKHGLDGLPSLTAFNC 434
QY 244 NIKQDINCR .....SFL .....AL 259
DB 435 IDYRAGLGKKYFSPRHISQEFYFVANSFVKSQDQVQVQVQVQVQVQVQVQVQV 494
QY 260 TSRLFGLLNEETKRSHEKSLCKWKSDELWQVLLQW:QSKAQSDGSLTQOQ SLEFFSL 318
DB 495 SLQELFVYKLTGCTLSLGLKFCFRIAFSVQQLNRP ---KEQ:EAIKYKSNWLEP--SL 549
QY 319 YEIQEEFFIQALSHFQVIVVSNIASMGMN 349
DB 550 YDSKIKLTCG-----IQMKDVLNVQHL 573

RESULT 9
Q8CAW5 PRELIMINARY; PRT: 925 AA.
AC Q8CAW5;
DT 01-MAR-2003 (TREMblrel: 23, Created)

```

```

DT 01-MAR-2003 (TREMblrel: 23, Last sequence update)
DT 01-MAR-2003 (TREMblrel: 23, Last annotation update)
CS Mus musculus (Mouse);
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
CX NCBI_TaxID:10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Ovary;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium.
RA the RIKEN Genome Exploration Research Group Phase 1 & 11 Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RJ Nature 420:563-573(2002);
DR EMBL; AK054264; BAC35710.1;
SQ SEQUENCE 925 AA; 94021 MW; 21D71DEABEAPCA31 CRC64;

Query March 19, 98; Score 443; DB 11; Length 825;
Best Local Similarity 24.5%; Pred. No. 4.5e-34;
Matches 130; Conservative 100; Mismatches 168; Indels 32; Gaps 15;

QY 2 NOSATECSMDLPSFCMPSPERSAPIQELRPVPERLFIIDGDELKSPHPQGWCLWEEKRPT 57
DB 239 NPSWPCCVIRDLIG .....KELFVVDGDEL--TF PAGAL:RDLQ 278
QY 58 LQWEKRPTELLNSLRKKLPELSLITTRPTALKRHLRHLHPRHPRHVEILGFSHAERK 117
DB 279 GDMTVKRPVEVLGSLKRRKAPHATLLVTRTQSLQIFVWQDQPLAVETLGLSESEKQ 338
QY 118 EYFYKFNAAE:-----QAGOVNRYVRNRP:FTWCVPVLYWV:TVLQGLCEHEDLQTS 168
DB 339 EYFOKYFEDEGESEDKSGKALRAKREVRNADLYQWASLPTACGIFCLCLELRKXKKE 398
QY 169 LLROTSRTTAVVYLYL:ISLQPKPGAFRLQPPNQ ---RGICSLAAGLWQ ..... 218
DB 339 ULSTLCQTYTKYKFNFP:CEVPSSETCEDHL-----NEEFQILFKKICILAAANSLLEQVEIL 454
QY 219 .....KILFEQC .....LRKHGLDGED 236
DB 455 CEDPFLK:KALNNLHFWVWRHLFKQSSSTHCL:SF:CLGIQQLAAAIFFVQLGQSKG 514
QY 237 VSAFNYK:IFKNDINCERSFLATSR:FLGLNNESTRSHLEKSLCKWKSVP:KNCILLQW: 296
DB 515 VSKYSIQ:KMSREARLQKPLPFGVGLNNE:TR:QELKATTFGCGQISTEVKPKF:SC- 573
QY 257 QSKAQSDAST:QGSSEIFFSCVIEIQEEFFIQALSHFQVIVVSNIASKVEHNVSSFLK 356
DB 574 ESGNKELLIMNQEL:LSGLYESQEGFVKRANVLFEDEL:HLKTS:DLIHASFLK 631
QY 357 KCRSAQVILNRYGATYSADGEDRARCAGAHLLVQ:---LPERTVLLDAYSEHLAA:--- 409
DB 642 KSNQ:QTM .....SLKVEKAVFPENVAALSTAKHQS:SPQFR 669
QY 410 .....IQ .....TNPNILIE:SLYANALGSRGKILCOGLRHPNCKLNL 445
DB 670 MLTFWTDKFDTFNSKNKLVFLDIESFNSSALSTREKLPASCCQKV 719

RESULT 10
Q8EPG7 PRELIMINARY; PRT: 657 AA.
AC Q8EPG7;
DT 01-MAR-2001 (TREMblrel: 16, Created)
DT 01-MAR-2001 (TREMblrel: 16, Last sequence update)
DT 01-MAR-2003 (TREMblrel: 23, Last annotation update)
DE ANI-like protein.
CS Mus musculus (Mouse).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
CX NCBI_TaxID:10090;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN=129/5vJ;
RX MEDLINE:213:0002; PubMed:11416212;
RA Lane R.P., Cutforth T., Young J., Athanasiou M., Friedman C.,
SA Bowen L., Evans G., Axel L., Hood L., Frank B.G.
RT "Genomic analysis of orthologous mouse and human olfactory receptor
RL loci."
RL Proc. Natl. Acad. Sci. U.S.A. 98:1199-1206; 2001.
CR EMBL: AF121233; AAC45189.1;
CR InterPro: IPR007097; IPR RN1n.
CR InterPro: IPR007111; NACHT NTPase
CR PROSITE: PS50503; LRR_R1;
CR PROSITE: PS50837; NACHT;
SC SEQUENCE 657 AA; 73410 MW; F4HCD4BPA1AAR? CR64;

Query Match 16.6%; Score 193.5; DB 11; Length 657;
Best Local Similarity 40.5%; Pred. No. 2, 2e-26;
Matches 104; Conservative 54; Mismatches 116; Indels 67; Gaps 9;

QY 161 QQLGEGGGLPCTSRITTAIVMYLISLNGPKP-RAPRQGFPPNQ---RQLGSLAAGLW 216
DB 161 QQLGEGGGLPCTSRITTAIVMYLISLNGPKP-RAPRQGFPPNQ---RQLGSLAAGLW 216
DB 161 QQLGEGGGLPCTSRITTAIVMYLISLNGPKP-RAPRQGFPPNQ---RQLGSLAAGLW 216
QY 217 NQKILFEQDLKRGLOGEDVSAPAKWLPKDI--NC--- 252
DB 217 NQKILFEQDLKRGLOGEDVSAPAKWLPKDI--NC--- 252
DB 217 NQKILFEQDLKRGLOGEDVSAPAKWLPKDI--NC--- 252
QY 218 TNKHVLYQQRKHGHELTREDEILLEDAKVLCQTHFENQVYTHLHVQGFPAALFYLLRE 277
DB 218 TNKHVLYQQRKHGHELTREDEILLEDAKVLCQTHFENQVYTHLHVQGFPAALFYLLRE 277
DB 218 TNKHVLYQQRKHGHELTREDEILLEDAKVLCQTHFENQVYTHLHVQGFPAALFYLLRE 277
QY 263 -----ERSFADTSE---EFGLNNEETRSLSKSLQWKVSPHI 289
DB 263 -----ERSFADTSE---EFGLNNEETRSLSKSLQWKVSPHI 289
DB 263 -----ERSFADTSE---EFGLNNEETRSLSKSLQWKVSPHI 289
QY 289 KNDLQWISQKASDSTGCGSLEFSTLYLHQLLEFLKALSHFQVIVNSIASKMEH 348
DB 289 KNDLQWISQKASDSTGCGSLEFSTLYLHQLLEFLKALSHFQVIVNSIASKMEH 348
DB 289 KNDLQWISQKASDSTGCGSLEFSTLYLHQLLEFLKALSHFQVIVNSIASKMEH 348
QY 338 REELALRELEKQDSSLSQRFQRLHCTVETQCEFTQALMYFKITV-RVDFEPQ 396
DB 338 REELALRELEKQDSSLSQRFQRLHCTVETQCEFTQALMYFKITV-RVDFEPQ 396
DB 338 REELALRELEKQDSSLSQRFQRLHCTVETQCEFTQALMYFKITV-RVDFEPQ 396
QY 349 MVSSFKLKRCSSACVLHLYCATYSAGDEDPAR CSACAHILVQCPERTVLDAVSEHLA 407
DB 349 MVSSFKLKRCSSACVLHLYCATYSAGDEDPAR CSACAHILVQCPERTVLDAVSEHLA 407
DB 349 MVSSFKLKRCSSACVLHLYCATYSAGDEDPAR CSACAHILVQCPERTVLDAVSEHLA 407
QY 397 KIVSFLKCHTCTKMLRPARADKMLDPAEMEDGAAVGVTH 446
DB 397 KIVSFLKCHTCTKMLRPARADKMLDPAEMEDGAAVGVTH 446
DB 397 KIVSFLKCHTCTKMLRPARADKMLDPAEMEDGAAVGVTH 446
QY 408 AALCTNPNLIELSYRNAGSGROVELLAQGRHPNCKLQD 448
DB 408 AALCTNPNLIELSYRNAGSGROVELLAQGRHPNCKLQD 448
DB 408 AALCTNPNLIELSYRNAGSGROVELLAQGRHPNCKLQD 448
QY 447 SVLRTNSLEIMELYBRLGSLKLNKLNELSHPKCKLQK 457
DB 447 SVLRTNSLEIMELYBRLGSLKLNKLNELSHPKCKLQK 457
DB 447 SVLRTNSLEIMELYBRLGSLKLNKLNELSHPKCKLQK 457

RESULT 1:
Q87NF5 PRELIMINARY; EST; 1020 AA.
ID Q8K32C
AC Q8K32C
DT 01-OCT-2002 (EMBL:rel. 22, Created)
DT 01-OCT-2002 (EMBL:rel. 22, Last sequence update)
DT 01-MAR-2003 (EMBL:rel. 23, Last annotation update)
DE "Cloning, sequencing and expression analysis of the murine Nod2/Card15 gene."
OS Mus musculus (Mouse).
GN CARD15.
NCBI_TaxID=10090.
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX 11
RN 11
RP SEQUENCE FROM N.A.
RC STRAIN=BAUS/C.
RA Iwanaga Y., Davey M.P., Martin T.M., Plack S.R., DePriest M.S.,
RA Baugh M.M., Suing C., Rosenbaum J.T.
RT "Cloning, sequencing and expression analysis of the murine Nod2/Card15 gene."
RT Submitted (JUN-2002) to the EMBL/GenBank/DBCC databases.
RC EMBL: AF542774; AAC16002.1;
DR MGD: MGI:2429197; Card15.
DR InterPro: IPR001315; CARD.
DR InterPro: IPR007091; LRR_RN1n.
DR InterPro: IPR007111; NACHT NTPase.
DR InterPro: IPR001310; TARD.
DR PROSITE: PS50409; CARD; 2.
DR PROSITE: PS50503; LRR_R1;
DR PROSITE: PS50837; NACHT;
DR PROSITE: PS01337; TARD; 1.
SQ SEQUENCE 1020 AA; 12356; MW: 25049056CF70FBB CRC64;

Query Match 10.6%; Score 453; DB 11; Length 1020;
Best Local Similarity 21.8%; Pred. No. 2e-15;
Matches 126; Conservative 79; Mismatches 208; Indels 166; Gaps 20;

```

```

QY 5 ATEGSMGLIFS -CWPEPSAP--LDELIVPERLLFIIDGFDELKPSHD - POS 54
DB 320 AKPLSUSTSFEHCWCDVACQVQFQLDHPORVLLTDLGLDEFFKPFTRCRHRHSPID 379
QY 55 PWKLCWEKRPTELLKSLKPKKLLPELSLITTRPTALEKLRLLRPHRPHVILIPSEA 114
DB 380 P -----TSVQTLAFNLQGNLLKNAKCVLSRPRVAVSALPKFVY TELQKJFSEE 430
QY 115 ERKEFYKYKHNAEACQVYVYDKRELFYKFPVPCVWVYCTCQQQLGGGLLRQTS 174
DB 431 GIOYLKRRHREFGVADRLCLIQATSAIRGCHLPVFSMMVSPCHRELLQN--RGFP 487
QY 175 RTTAVVYML-----YLGLSXGPKPCAPPLQPPPGGSGS-----LAADGLWQKILFER 224
DB 488 TTSTDMYILLIQHFLHASPDSFGLGPGGLGSRJSTLILHGHJALRGLAMSCVYFSA 547
QY 225 QDLKXKGLDGEDVS-----AFINYNK--FQKDNCRSEFLALT 260
DB 548 QGLCAAVDSDSLGLFVRAQSSVPSKAFLEFLHITFCCHFAFVLAVSATSVASLK 657
QY 261 SRLEFGLNSETSHLEKSLC-----WKVSNH-----DQST-----LQGSLEFFSCLY 287
DB 608 HLESCGRGLSSLLGRLLFNLCIQSPYKKGSEALCKAEPHNQITAAFLAGLLSQHR 667
QY 288 -----IKWLLQWIOSKACS-----DQST-----LQGSLEFFSCLY 319
DB 668 DLAAQCVSRVLLQORARSCIAHSLRHFHSIFPAVIGEAKSMHAPGFWLIRSEY 727
QY 720 EIGEEEFQALSHFOVIVVSNIAKMKHVMVSSFC-----LKKRCSAQVHL 366
DB 728 EMQFEQLQAEAVRLD-----CHLKLITFRVCPACBAAALAFVCHLQRFVALQL 777
QY 367 YGATYSAGGDRAP-----GSA-----GANTIL VQRPETV-----LED 480
DB 778 DYNEVGPVGVSEQLRFLGVCTALYDKNNDIDRGNATQVFACTQEQLOKLAENKRLD 837
QY 401 AYSHLAAALCTNELLISLYNAKDSGVKLLAQNR 439
DB 838 ACAGSMAKLHAHKONFLSLRVGRNHITAAAEVLAQNLK #76

RESULT 13
Q8BUT6 PRELIMINARY: PRT: 978 AA.
AC Q8BUT6;
DI 01-OCT-2002 (TrEMBLrel. 22, Created);
DI 01-OCT-2002 (TrEMBLrel. 22, Last sequence update);
DI 01-MAR-2003 (TrEMBLrel. 23, Last annotation update);
DE FLOC3348 protein (fragment);
GN FLOC3348;
OS Homo sapiens (Human);
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo;
OX NCBI_TaxID=9606;
RN 1;
RS SEQUENCE FROM N.A.
RC Tissue-Spleen;
RA Jikuya H., Takano J., Kikano J., Nagase T., Ohara O.;
RC "The nucleotide sequence of a long cDNA clone isolated from human
RC spleen.";
BL Submitted (JUL-2002) to the EMBL/GenBank/UDC databases.
DR EMBL; AKC90431; BAC03412;
DR InterPro; IPR001395; AldoKet_red;
DR InterPro; IPR005566; Lipocin_CyFABP;
DR InterPro; IPR007091; LRR_RNinh;
DR InterPro; IPR003590; LRR_RNinh_sup;
DR InterPro; IPR007111; NACT_NTPase;
DR SMART; SMC3368; LRR_R1; 2;
DR PROSITE; PS00063; ALDOCKETO_REDUCTASE_3; 1;
DR PROSITE; PS00213; LIPOCALIN; 1;
DR PROSITE; PS05053; LRR_R1; 1;
DR PROSITE; PS50837; NACT; 1;
FT NCTER 1 1

```

```

SQ SEQUENCE 778 AA: 95297 MW: ALEMAC43FF32836C CRC64;
Query Match 10.48; Score 246.5; DB 4; Length 778;
Best Local Similarity 23.28; Pred. No. 6.1e-15;
Matches 132; Conservative 81; Mismatch 190; Indels 167; Gaps 23;

QY 13 LIPSCWP--ERSAFICR--RYPERLLFIIDGFDELKPSFHPDQGPWLCWEKEKPTELL 69
DB 219 LICSFPHRVEBSL-----AVAYPAEALLLDGDECRTPLDFTNTVACTDPKKEIPVDH 265
QY 70 INSLTRKKCLPELSLITTRPTALEKLRLLRPHRPHVILIPSEA 114
DB 266 ITNIRGNLFFAVSNTSRPSASSQPGCLV-DRMTLRGNFEEELKVCLEQMPPE-DQ 323
QY 130 A -GOVTVVRKRELFYKFPVPCVWVYCTCQQQLGGGLLRQTS-----RTT 177
DB 324 ALLGMLSSQVADRALYACTVPAFCRLTGML-----GHWRSRITGPDAAELWPPRL 177
QY 178 TAVYMLYLLSLM-----QPKCAPPLQPPN-----QRLCSJAAGGLWNQILFEEQDL 227
DB 378 CELYSWYFPMALSGEQEKGKASPRIQVAHGSRKMVSTGLCAFHGLLKKKYVFEQDM 437
QY 228 RXHGLD-----CEDVSAPLN-----GHWRSRITGPDAAELWPPRL 246
DB 438 KAFGVDAILOQAPISCFQRETLASSVAYCFTHLSQEFVAAAYVYGASRAIPLDLET 497
QY 247 -----QKQINCRSEFLALTSHRFLQGLNNEETPSHLEKSLKWKVSPH 267
DB 498 EESGVMPRLGFLTHFRSAAQRAMQAEQGRDLDVFLAFUSGLLSPRYKALLAGSL----- 550
QY 288 IKWLLQWIOSKACSQSGSTLQGS-----SLFFSCLYEICEEF-----IQQAL-- 331
DB 551 --LAQGEHQAYRTOVAELQGLRQDAACARAINVLHCLJHELCHTELARSEVSEAKES 606
QY 332 -----SH--FOVIVVSNIAKMKHVMVSSFC-----LKKRCSAQ----- 362
DB 607 GALARLGTGAHRAALAYLIQVSDAAGEANLSLSQGVLOSLLPCLLYCRKLRLDTQCF 666
QY 363 --VHLVQATYSAGGDRAPGSAQANTLLVQLPRTVLLDAYSEHLAAALCTNELLIED 419
DB 667 QPFWMEDGLSVLS--GKT--CHLKLITFRVCPACBAAALAFVCHLQRFVALQL 716
QY 420 SLPYNAKDSGVKLLAQNRHNPCKLQNR 449
DB 717 DKGNSHFGAKALADALE INRTLTERR 745

RESULT 14
Q8BUT6 PRELIMINARY: PRT: 706 AA.
AC Q8BUT6;
DI 01-MAR-2003 (TrEMBLrel. 23, Created);
DI 01-MAR-2003 (TrEMBLrel. 23, Last sequence update);
DI 01-MAR-2003 (TrEMBLrel. 23, Last annotation update);
DE Similar to CASPASE RECRUITMENT domain protein 4 (fragment);
OS Mus musculus (Mouse);
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus;
OX NCBI_TaxID=10090;
RN 1;
RS SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
RC MEDLINE=22354683; PubMed=13466851;
RA The FANTOM Consortium;
RA the RIKEN Genome Expression Research Group Phase I & II Team;
RC "Analysis of the mouse transcriptome based on functional annotation of
RC 60,770 full-length cDNAs.";
DR EMBL; AKC82663; BAC38566;
FT NCTER 1 1
SQ SEQUENCE 706 AA: 29523 MW: 66814F891C997DF5 CRC64;
Query Match 9.84; Score 233; DB 11; Length 706;

```

